

401
8970

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by

L. B. BIRCH

Associate Editors:

P. E. VERNON and DORIS M. LEE

Advised by

SIR CYRIL BURT
R. A. C. OLIVER
E. A. PEEL

SIR FRED J. SCHONELL
S. WISEMAN

Vol. XXXV

1965



Published by

METHUEN AND CO., LTD., 36, ESSEX STREET, STRAND, LONDON, W.C.2.,
for the British Journal of Educational Psychology, Ltd.

Registered Offices: 42, Hill Street, Birmingham, 5. This address should *not* be used for
letters or books intended for the Editor, Authors, or the Publishers.

On behalf of

The British Psychological Society and the Association of Teachers in Colleges and Departments
of Education.

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS, LTD., 42, HILL STREET, BIRMINGHAM, 5.

Buy Edm. Research
Date 8.9.70
Pasted 8-409
Index No.

THE EFFECT OF PRACTICE ON PERFORMANCE IN A READING ATTAINMENT TEST

By C. NETLEY, S. RACHMAN AND R. K. TURNER

(Institute of Psychiatry, Maudsley Hospital, Denmark Hill, S.E.5, London University)

SUMMARY. Although it is highly desirable that educational progress should be assessed by means of standardised attainment tests, this procedure often encounters difficulties. The artificial inflation of re-test results by practice effects is a particularly awkward problem. The present study was designed to measure the extent of this effect on parallel forms of a recently developed attainment test, namely the Neale Analysis of Reading Ability. It was found that this Test is of high reliability and is insensitive to practice effects.

I.—INTRODUCTION.

THE present investigation arises from difficulties encountered in evaluating progress made by backward readers who have been given remedial coaching. The increases in scores on attainment tests that occur on retest are usually interpreted as 'true' gain due to the coaching. However, such increases may also be attributed to the practice effects on the attainment test used. Increases which are 'built into' the test itself would be expected to occur in the absence of specific coaching, and there is evidence to suggest that they may be in excess of what would be expected assuming normal maturational changes.

In a carefully designed study, Curr and Gourlay (1960) showed that when 9-10 year old children were retested after various time intervals on reading attainment tests, highly significant gains occurred in the absence of specific coaching. These practice effects were highest on tests of reading comprehension (26.9 months and 12.2 months for average and below average readers, respectively, after a six-month interval). The gains on mechanical reading were slightly lower (14.1 and 11.8 months after six months, for the average and below average groups). After a three months time interval, the gains in reading comprehension for the average and below average groups were 18.2 and 10.4 months, respectively, whilst for mechanical reading they were 8.0 and 6.2 months. As can be seen from the above figures, the below average readers showed consistently lower gains. Significant differences both in initial scores and in practice effects between the four schools sampled were also found.

The aims of the present study were to obtain data on the test-retest reliability of a reading attainment test and to determine the extent of practice effects occurring between tests in the absence of specific coaching. As it was our intention to use the data in the evaluation of coaching given to backward readers attending the Children's Department Clinics at the Maudsley Hospital, it was decided to test a sample of children drawn from a school in the Hospital area.

II.—THE INVESTIGATION.

Ninety-eight children in the three oldest classes (unstreamed) of a local primary school were allocated to four test groups—AAA, ABC, BA and AB, by order of name on alphabetical class lists. Owing to very high absence rates due to illness, thirty children were excluded from the trial (most of the testing was carried out during a period of severe weather). The remaining sixty-eight

children were all tested individually. The mean ages of the groups are presented in Table 1.

TABLE 1
THE AGE COMPOSITION OF THE SAMPLE.

	No. of Children	Mean Age	Standard Deviation
Group ABC	17	9.10 years	1.18
Group AAA	16	9.27 years	1.18
Group BA	17	9.85 years	1.06
Group AB	18	9.43 years	1.22

Tests between the mean ages of the four groups revealed no significant differences.

III.—THE TEST.

The Neale Analysis of Reading Ability (1963) was selected for study since it offers the following advantages: (a) the provision of parallel forms, and (b) three separate scores (reading accuracy, comprehension and rate) are obtained in a single test administration.

The Neale Test comprises six short stories providing a continuous reading scale from the age of 6–12 years. The child continues to read each story in turn until he reaches one in which a criterion number of errors is made. At the end of each passage comprehension questions are asked. A reading rate score is obtained on the basis of the average number of words read per minute.

Adequate standardisation data on over 2,000 children is available, and information about the test's reliability, validity and parallel forms is presented in the Manual.

IV.—TEST ADMINISTRATION.

As a retest was to be made, the standard instructions given in the Manual (1963) were followed with one difference. When errors were made the correct answer was not provided*; the child was encouraged to try and then told to continue with the next word.

The testing was carried out by post-graduate members of the course for the Post-graduate Diploma in Abnormal Psychology and three staff members, all of whom were at the time working in the Children's Department. Test conditions were not ideal owing to a shortage of room in the school. However, it was considered unlikely that the validity of the results was significantly impaired.

V.—DESIGN OF THE STUDY.

Retests were made over two time intervals: (a) one week and (b) fourteen–fifteen weeks. It was assumed that practice effects due to increased familiarity with the test would be maximised during the first period, whilst maturational factors could be assumed to be relatively more evident over a fourteen-week period.

It was decided to retest on the same and parallel forms, so as to afford a comparison of the practice effects occurring not only when the children were given the same form of the test on all occasions, but also when different forms

* Owing to a misunderstanding, some children were accidentally given the correct answers.

were used (i.e., forms A, B and C in order). In addition, two of the four test groups were given the same forms (A and B) but in different orders and over different time intervals. This was done in order to determine the relationship between inter-test time intervals and score changes.

TABLE 2
ORDER OF TEST ADMINISTRATION.

	Initial Test	Retest at 1 week after initial test	Retest at 15 weeks after initial test
Group ABC	A	B	C
Group AAA	A	A	A
Group BA	B	A	—
Group AB	A	—	B

(The letters indicate the three forms of the Neale Test.)

Thus, a child in group 2 was tested on form A on all three test occasions, whilst a child in group 4 was given form B and retested after one week on form A, but was not given a third retest.

VI.—RESULTS.

As a first step in the analysis of the test results, the means and standard deviations for all groups were computed. These are shown in Tables 3, 4, 5 and 6. All test results which were in the form of years and months were converted to years and fractions of years.

TABLE 3
MEANS AND STANDARD DEVIATIONS OF THE TEST RESULTS IN YEARS
FOR GROUP ABC (n=17).

	Accuracy	Comprehension	Rate
Form A \bar{x}	8.24	7.90	8.81
s.d.	1.39	1.40	2.15
Form B \bar{x}	8.34	8.10	8.95
s.d.	1.46	1.50	1.88
Form C \bar{x}	8.36	8.32	8.70
s.d.	1.47	1.55	2.34

TABLE 4
MEANS AND STANDARD DEVIATIONS OF THE TEST RESULTS IN YEARS
FOR GROUP AAA (n=16).

	Accuracy	Comprehension	Rate
Form A (1st Occasion) \bar{x}	8.71	8.61	8.89
s.d.	1.77	1.96	2.11
Form A (2nd Occasion) \bar{x}	8.89	8.89	9.26
s.d.	1.76	1.82	2.01
Form A (3rd occasion) \bar{x}	9.29	9.56	9.40
s.d.	1.74	2.02	2.30

TABLE 5
MEANS AND STANDARD DEVIATIONS OF THE TEST RESULTS IN YEARS
FOR GROUP BA ($n=17$).

	Accuracy	Comprehension	Rate
Form B \bar{x}	9.15	9.49	9.48
s.d.	1.60	1.86	2.17
Form A \bar{x}	9.24	9.24	9.30
s.d.	1.52	1.86	1.91

TABLE 6
MEANS AND STANDARD DEVIATIONS OF THE TEST RESULTS IN YEARS
FOR GROUP AB ($n=18$).

	Accuracy	Comprehension	Rate
Form A \bar{x}	8.01	8.01	8.25
s.d.	1.41	1.47	1.61
Form B \bar{x}	8.27	8.60	8.72
s.d.	1.44	1.83	1.99

These results indicate that the children's reading ages were somewhat behind their chronological ages.

Further analysis of the data is presented below under separate headings :

Reliability. As one of the principal aims of the study was to determine the retest reliability of the Neale Test, product-moment correlations were computed between pairs of test results for each testing occasion with form A. The results of these calculations are shown in Table 7. All correlation coefficients were found to be statistically significant ($p < .01$) and indicate that all of the Neale subtests possess a satisfactory level of retest consistency. There is a suggestion in the data (most pronounced in the case of the Rate results) that the time interval separating pairs of test occasions is a relatively unimportant consideration in determining the size of the association between them. Rather, it appears that the greater the number of prior testings on Form A, the higher the resulting correlation coefficient between any pair.

While all the correlation coefficients were significant, the lower r for rate between the first and second testings suggests that any conclusions concerning dramatic changes in rate scores should be considered in relation to the correlation coefficient.

TABLE 7
RE-TEST CORRELATIONS WITH VARYING TIME INTERVALS WITHIN GROUP AAA.

	A_1-A_2 (1 week)	A_1-A_3 (15 weeks)	A_2-A_3 (14 weeks)
Comp.915	.868	.942
Accur.915	.941	.953
Rate775	.814	.913

Intercorrelation between subtests. In Neale's published study (Neale, 1963) there are no data indicating the relative associations between the Test's three subtests even though such information would be useful in interpreting subtest discrepancies. The intercorrelations of the Accuracy, Comprehension and Rate subtests for Forms A and B are given in Table 8. It can be seen that they are all quite high and, indeed, all are statistically significant ($p < .01$).

The lower correlation between rate and the remaining two subtests may possibly be due to the relatively lower reliability of the Rate scores.

TABLE 8

INTERCORRELATIONS OF ACCURACY, COMPREHENSION AND RATE SCORES ON INITIAL TESTING WITH FORMS A AND B.

	Accuracy	Comprehension
Form A (n=51):		
Comprehension ..	.903	—
Rate752	.685
Form B (n=17):		
Comprehension849	—
Rate842	.727

Comparison of Forms. There is little information available on the relative difficulty of the various forms of the Neale Test. The assumption has been made that Forms A, B and C are parallel and may be used interchangeably.

In an attempt to confirm this assumption, the results for all subjects tested initially with Form A were compared with the results of subjects tested initially on Form B. These results are summarized in Table 9 and suggest that Form B is somewhat easier than Form A. The difference between Forms A and B reaches statistical significance in the case of Comprehension and approaches significance for Accuracy and Rate.

TABLE 9

TESTS OF DIFFERENCES BETWEEN FIRST TESTING ON FORMS A AND B.

	N _A	N _B	\bar{A}	\bar{B}	S.E. A-B	\pm	P *
Accuracy	51	17	8.31	9.15	.427	1.97	>.10 <.05
Comprehension	51	17	8.16	9.49	.465	2.87	>.01
Rate	51	17	8.64	9.48	.555	1.52	>.20 <.10

* 2 tailed test.

It is not possible, however, from these data or from the data in the Manual, to conclude firmly that Forms A and B are of unequal difficulty. While subjects were assigned to their experimental groups in a random manner, it appears

that the BA group whose results were used in the above comparison were, on the average, slightly (but not significantly) older than subjects tested initially on Form A (see Table 1). In addition, it is not possible with Neale's data to ascertain whether testing orders were controlled. However, by combining the results obtained from the BA and AB groups, it is possible to produce more conclusive data relating to the relative difficulty of Forms A and B. The results of correlated *t* tests for these data revealed no significant differences. It may be concluded that Forms A and B are not reliably different in their average results. It appears likely that the significant difference in Table 8 is due to the uncontrolled age variable inherent in this comparison.

In an effort to gain further information about the presumed equivalence of Forms A and B, product moment correlations were computed between the combined results of groups AB and BA. Table 10 shows the results of these calculations. All correlations are highly significant ($p < .01$) and indicate that Forms A and B are highly related.

TABLE 10

PRODUCT MOMENT CORRELATION COEFFICIENTS BETWEEN FORMS A AND B FOR THE COMBINED RESULTS OF GROUPS AB AND BA.

Accuracy913	n=34
Comprehension929	n=34
Rate812	n=34

The results of both correlations and tests between means indicate that Forms A and B are probably parallel. A possible exception to this conclusion is that Form B's Comprehension subtest may be somewhat easier than the Form A measure.

Practice Effects. One of the main aims of the present study was to determine what changes, if any, would be expected as a result of repeated testings on the Neale Test. It is obvious that such information is useful if the test is to be used in educational and clinical settings where changes in test scores must be evaluated.

Within the present design two factors could be expected to produce increases in test scores as a result of repeated testings. The first might be termed score increase as a result of the educational process, i.e., the children were receiving their usual tuition from school after their initial testing. Secondly, practice effects on a particular test form or experience with the testing situation could lead to score increases. In an effort to separate these factors, two analyses of all the data were carried out. They were:

- (1) *Analysis of Variance.* This would reveal any changes in test scores due to the combined effect of educational experience and practice.
- (2) *Analysis of Covariance.* Using this statistic and employing the subject's age at each testing, which may be assumed to reflect the amount of formal educational experience of each child, as the predict or variable changes due to practice alone may be evaluated.

GROUP AAA.

Analyses of variance and covariance of all subtests for this group were carried out. No comparison achieved significance ; this shows that no reliable changes in test scores occur as a result of three testings on Form A over a three-month period. This rather surprising result indicates that, at least with Form A, neither practice nor the educational process lead to any improvement in test score. It is possible that three months is too brief an interval to expect reliable changes on this test. This may be particularly relevant when dealing with a somewhat backward group. It should be noted that the means of the three groups do increase with repeated testings suggesting the existence of a trend.

GROUP ABC.

When this study was designed it was expected that significant improvements in test scores would occur as a result of repeated testings on the same test form. Because of this expectation, group ABC was tested in order to determine whether score changes could be expected as a result of a repeated acquaintance with the test situation itself, i.e., do changes occur which are not directly attributable to practice on a particular form of the test.

Analyses of variance and covariance were carried out and again, no significant changes were found. In view of the absence of reliable changes in Group AAA, this is not unexpected.

GROUP AB (1 week) and GROUP AB (3 months).

Analyses of these two groups were carried out to determine whether the time interval between testings was related to the magnitude of test score change. The results were not significant in either case. The absence of significant score changes in either group indicates that, within the time range examined, changes in test scores are not differentially related to the test retest interval. The absence of such changes suggests that the time intervals used in this study were too brief.

GROUP BA (1 week).

This group was included primarily in order to have counterbalanced data to combine with group AB (1 week). Variance and covariance analyses were carried out ; no significant changes occurred.

VII.—DISCUSSION AND CONCLUSIONS.

The test-retest reliability of the Neale Test is satisfactorily high, particularly on the comprehension and accuracy measures. The relatively lower reliability of the Reading Rate measure may account for the lower correlation between this measure and comprehension and accuracy. The latter two measures are highly correlated.

There is a suggestion that Forms A and B although correlated, are not entirely equivalent. Our data do not permit a firm conclusion on this matter, but the indications are that Form A may be more difficult than Form B in one respect at least (comprehension). As Form A has a high reliability and is not distorted by practice effects, it would seem that the exclusive use of this form of the test can be recommended (particularly when the child is to be retested).

In view of the findings of Curr and Gourlay (1953, 1960) on this subject, we paid particular attention to practice effects. Unlike the tests which they used, scores on the Neale Test appear to be unaffected by repeated exposures. No significant increases were obtained when the test was re-administered after one

week. The scores did show some increase after a period of three months, but did not reach statistical significance. It can be argued, of course, that this result shows the test to be insensitive to real gains. We feel, however, that three months is perhaps too short a period in which to observe genuine improvements in reading ability, particularly in a group of children who were backward in reading. It should be remembered, however, that a larger practice effect is obtained from bright children than from dull children (Curr and Gourlay, 1960). Having established that the test is not inflated by practice effects, it will now be possible to ascertain its sensitivity over a longer period, e.g., one to two years.

A more general conclusion can also be drawn from this enquiry. Not all reading tests are subject to retest distortions and they can be usefully applied in evaluation studies. Curr and Gourlay's discouraging results are, nevertheless, of some importance and indicate the need to ensure that any test which is used for evaluation purposes should be carefully assessed. Their work also underlines the need to include adequate control groups in follow-up studies.

VIII.—REFERENCES.

- NEALE, M. D. (1963). *Neale Analysis of Reading Ability Manual*. London : Macmillan and Co.
- NEALE, M. D. (1956). *The Construction and Standardization of a Diagnostic Reading Test in a Survey of the Reading Attainment and Attitudes of Primary School Children*. Unpublished Ph.D. Thesis, University of Birmingham.
- CURR, W., and GOURLAY, N. (1953). An experimental evaluation of remedial education. *Brit. J. Educ. Psychol.*, 23, 45-55.
- CURR, W., and GOURLAY, N. (1960). The effect of practice on performance in scholastic tests. *Brit. J. Educ. Psychol.*, 30, 155-167.

(Manuscript received 1st May, 1964)

ENVIRONMENTAL HANDICAPS AND INTELLECTUAL DEVELOPMENT*: PART I

By PHILIP E. VERNON

(*University of London, Institute of Education*)

SUMMARY. This is the first of two articles whose object is to throw light on the development of, and retardation in, abilities by comparison of patterns of test scores with assessments of environment, both within and between certain contrasted cultural groups.

A series of verbal tests is described, some group, some individual, which were applied to a representative group of 100 boys close to 11:0 years in South-East England, and to a sample of fifty similar boys in the West Indies. These included tests of educational attainments, memorizing, vocabulary, and a battery of items based on Piaget's investigations of concept development. The latter was shown to be factorially complex. Mean West Indian scores varied widely, ranging from an English deviation quotient of 72 on Vocabulary to 94 on Spelling.

I.—INTRODUCTION.

Most psychologists nowadays would agree that it is unprofitable to talk about, or investigate, racial or ethnic differences in intelligence—for two main reasons. First, that the intelligence we can observe or measure is always the resultant of interaction between genetic potentialities and environmental pressures; and secondly, that intelligence is no one thing, but rather a name for a group of overlapping mental skills whose content depends considerably on what a particular culture values, or on what psychologists who belong to that culture like to include within their concept (cf. Vernon, 1960). It is, however, much more meaningful to study the effects of controllable environmental factors on the development of various types of ability, either within a particular culture or between contrasted cultures.

Now Western European and North American psychologists have isolated a number of skills of proved importance for educational and vocational advancement which can either be classified, as Thurstone does, under the headings of such factors as Verbal, Number, Space, Induction and so forth or, according to English practice, as *g* plus a series of smaller group factors (Vernon, 1961a). In the developing nations of Africa, Asia and elsewhere, intellectual development often follows very different lines, depending on the cultural ideals and child-rearing practices and educational systems, so that inevitably members of these groups tend to perform relatively poorly on Western tests. But if these new nations are aiming at a Western-type technological civilization, as many of them are, it is essential that they should produce more individuals with Western type mental skills—more professionals, teachers, executives and technicians. What then are the main handicaps that lead to retardation in such abilities: physical factors of health and nutrition, simple economic limitations, lack of cultural stimulus in the home, perceptual deprivation, superstitious beliefs, child-rearing practices which discourage rather than encourage intellectual development, linguistic difficulties, or deficiencies in schooling? To provide definite answers in so complex an area of multiple causation would clearly

* Based on a paper read at the British Psychological Society, Annual General Meeting, Leicester, April, 1964.

require immense resources for longitudinal studies of children in controlled environments. But in the present writer's view, a useful start can be made by surveying the patterns or profiles of abilities on a rather wide variety of tests in different cultural groups, and by analysing any associations between peaks or troughs in abilities and some of the environmental factors suggested by previous research as most significant in intellectual development.

This article reports the results of a survey of a small sample in the West Indies, comparing it with a more representative English sample. It is planned to carry out similar surveys of Canadian Eskimo and Indian boys, of certain African groups, and of interesting sub-cultures in the U.K., such as Gaelic children, delinquent boys, immigrants, etc. In order to limit the scope to some extent, the investigations are confined to boys aged $10\frac{1}{2}$ —11 years in cultures where standard English is fairly well understood.

II.—SAMPLES AND PROCEDURE.

In the writer's view, it is better to test smallish samples with individual tests, so as to ensure that the tasks are thoroughly understood, than to conduct larger studies with group tests. An attempt was made to secure a reasonably representative South-East England group by testing thirty boys in an older London borough, fifty in two Hertfordshire country towns, and twenty in small villages, 100 in all. At a first visit to each of nine schools, a series of group tests was applied by the writer to all boys within the required age range. Norms were available for three of these tests—an Abstraction test, Vernon's Graded Arithmetic-Mathematics test and Mill Hill Junior Vocabulary Form A (Synonyms). From the 149 boys so tested, 100 were selected whose scores accorded closely with the expected means and distributions. In fact, the final sample was somewhat above the norms on all tests: thus, the median score in Arithmetic was $27\frac{1}{2}$ compared with a published norm of 25. But this is probably to be expected in South-East England where both bad slums and primitive rural conditions have largely disappeared, and where almost all pupils are more test-sophisticated than in the past. For example, the villages are occupied more by car-workers in factories fifteen miles away than by agricultural labourers, and even in the poorest London school it was rare to see a raggedly dressed boy, or one who did not get 6d. a day or more in pocket money. It may be that the welfare state will necessitate an upgrading in the norms for all the commonly used psychological tests (cf. Pilliner, *et al.*, 1960).

The selected boys were given a series of individual tests and one further session of group tests, by Mrs. D. Vernon, assisted by the writer—that is, some four hours of testing in all. It was hoped to apply a structured interview for eliciting standard information as a basis for assessing eight environmental variables considered of major importance in the development of abilities. However, in view of the natural caution of Heads and education officers, as much of this information as possible was collected from the Heads and/or class teachers, and the boys were questioned informally only on such 'safe' topics as leisure-time activities, including household chores, books and cultural pursuits, vocational aspirations, etc. Thus, the ratings, apart from A, C and D, are likely to be of low reliability. The eight variables were as follows:

A.—Regularity of schooling from 5 years to the present *versus* frequent changes, absences due to illness, or truancy.

B.*—Unbroken home (nuclear family) *versus* step-parent, desertion, father frequently absent, mother at work, grand-parents or others taking over parental functions.

C.—Economic status : father's (or substitute's) job, with a plus adjustment for earnings by other members of the family, and a minus adjustment for overcrowded home.

D.—Cultural stimulus : parents' or sibs' secondary education ; books and periodicals in the home, use of library ; visits to places of interest ; co-operation with the school and help with homework.

E.—Encouragement of initiative as shown by adventurousness of boy's pursuits, responsibility and maturity expected *versus* over-protectiveness or arbitrary subjection.

F.—Male dominance, identification with father (or male substitute) *versus* stronger influence of the mother.

G.—Family planfulness : rational and democratic home atmosphere and upbringing *versus* impulsive, emotional, inconsistent, feckless. (Obviously, this rating was particularly impressionistic.)

H.—Linguistic background : encouragement of good speech *versus* local dialect or slang, making for difficulties in communication at school.

It might be thought essential to evaluate such factors on the basis of parental interviews and home observations. But, apart from expense and time, it would be extremely difficult to obtain the co-operation of parents of normal English children. Moreover, the writer would hold that the details of training in early infancy which have been stressed in such researches as those of Sears, Maccoby and Levin (1957) and others are less important psychologically than the general climate and values of the home throughout childhood and up to the present time (Vernon, 1963), which can be judged fairly effectively by questioning a 10—11-year-old boy and his teachers.

The same procedure (with but minor alterations) had previously been applied to fifty boys in West Indian primary schools, with the helpful collaboration of the Institute and the Department of Education of the University of the West Indies. All the interviewing was conducted by one of two West Indian psychologists conversant with the local environment and with the mother-tongue or Creole ; and such testing as required detailed oral instructions was undertaken by senior students from the Department. The Piaget battery was given by D. Vernon, who became reasonably fluent in the dialect, and the present writer was responsible for performance tests requiring minimal instructions. Testing was often done under unusual conditions, for example, in the playgrounds, and there were many difficulties through absences, noises of chanting or corporal punishment in neighbouring classrooms, and so forth. The boys were drawn in groups of ten from five contrasted primary schools, chosen to represent the following sub-cultures : better-class urban, poorer-class urban, country town, sugar-cane plantation, and isolated rural small-holding. This group would be superior to the average for the whole population, since the economy of the island is predominantly agricultural, and a considerable proportion of boys of 10 upwards (usually, of course, the least able) attend school very irregularly. However, it would be reasonably representative of those male pupils who complete the elementary course and sit for the secondary school selection examinations (cf. Vernon, 1961b ; Manley, 1963).

* In the West Indian group, Stability *versus* Frequent Shifts from One Family or Place to Another, was assessed as an additional variable.

III.—THE VERBAL TESTS AND THEIR RESULTS.

Table 1 lists the 90th, 50th and 10th percentile scores for each test in the English and West Indian groups; non-verbal tests are described in a second article. In addition, the West Indian percentiles are expressed as quotients on English norms, with a standard deviation of 15. For example, the WI median of $15\frac{1}{2}$ in Arithmetic falls at the English 14th percentile, or 1.08 sigma units below the mean, which corresponds to a deviation quotient of 84.

Arithmetic. Vernon's Graded Arithmetic-Mathematics test. The first forty items were given with a twenty-minute time limit. The deficit on this test, represented by a median West Indian quotient of 84, is less serious than on several other tests, presumably because much time is devoted in the schools to drilling in mechanical arithmetic. Indeed, compared with the published norms (which may nowadays be too lenient in Britain) the West Indian boys are only $1\frac{1}{2}$ years backward, while their median school life is at least $1\frac{1}{2}$ years shorter than that of the top class in an English junior school.

English. An English attainment test, similar to current N.F.E.R. tests, was constructed by L. H. E. Reid (1961) for use in the West Indies. Four of the six sub-tests were employed: Reading Comprehension (5 minutes), Spelling (6), Parts of Speech (6), and Grammatical Usage (5), seventy items in all. The median quotient was 82, but on the Spelling sub-test, which measures an ability that can be highly drilled, the median was 94.

Vocabulary. Mill Hill Junior Form A, thirty-three multiple-choice items. In this and the English test, several sample items were written on the black-board and explained at length, in accordance with the procedure recommended by Schwarz (1961), so as to reduce the West Indian pupils' handicap of unfamiliarity with objective tests. A time limit of about 10 minutes was found sufficient.

Form B was also given individually, the pupil supplying his own definitions in his own time. Scores were so low that a change was made to Terman-Merrill

TABLE 1

COMPARATIVE PERFORMANCE OF ENGLISH AND WEST INDIAN BOYS ON VERBAL TESTS.

Test	English Percentiles			West Indian Percentiles			West Indian Deviation Quotients		
	90	50	10	90	50	10	90	50	10
Arithmetic	35	$27\frac{1}{2}$	14	29	$15\frac{1}{2}$	2	102	84	<65
Spelling	15	7	1	13	5	0	112	94	75
English total	63	48	13	44	$20\frac{1}{2}$	2	103	82	<65
Vocabulary (Mill Hill group test)	21	17	6	$15\frac{1}{2}$	8	2	104	83	71
Vocabulary (Terman-Merrill individual)	$21\frac{1}{2}$	15	10	12	8	5	87	72	<65
Word Memory ..	44	$30\frac{1}{2}$	$17\frac{1}{2}$	40	25	7	115	91	<65
Information Memory	28	22	14	23	10	2	104	72	<65
Concept Formation	5.25	3.5	1.4	3.7	2.65	1.6	101	90	82
Piaget battery (total errors) ..	$3\frac{1}{2}$	$9\frac{1}{2}$	16	7	$14\frac{1}{2}$	$23\frac{1}{2}$	107	86	75

Form LM Vocabulary when the English sample was collected. The percentile scores were 9-6-3, and according to Fraser Roberts and Dunson (1955), these are equivalent to 12-8-5 on the Terman-Merrill, or to Vocabulary Ages of 9 years, 7 years and 5 years, respectively. The median quotient relative to the English sample results was also very low, namely 72. Clearly, the West Indian pupil's lack of familiarity with standard English is one of his main handicaps, though he does acquire a fair degree of facility with the kind of reading and word usage involved in English tests and so performs considerably better on multiple-choice vocabulary items.

Word Memory. The next two tests were included to investigate the success of the drilling and rote learning for which West Indian education is notorious. They were adapted from the group tests described by MacKay and Vernon (1963) in a study of learning ability. After preliminary practice in learning a list of six monosyllabic verbs, pupils were given a list of twenty verbs, which they studied for five minutes, and then wrote out as many as possible. Another two minutes' study was given and another reproduction. The score was 1 for each word, regardless of spelling, plus $\frac{1}{2}$ for each one in correct rank order. Thus, the scores in Table 1 are out of total of 60, and the West Indian median quotient of 91 does indicate a well-developed capacity at this kind of task.

Information Memory. The result was much less favourable when meaningful information had to be learnt, namely, a median quotient of 72. After a practice series, the tester read out slowly a list of fifteen 'facts,' emphasizing the critical words, e.g.:

A barrel of beer holds 36 gallons—36 gallons.

Thereafter, a question was asked on each 'fact,' in a different order, e.g.:

How many gallons are there in a barrel of beer?

A further reading and set of questions followed. The median West Indian score was only 10/30, corresponding to a quotient of 72, despite the test being administered by a West Indian teacher.

Concept Formation (individual test). Twenty common objects were spread out at random, namely: a pencil, cigarette, box of matches, bar of chocolate, wooden bear, plastic cow, woollen tiger, orange, leaf, twig, artificial flower, nail-scissors, watch, toy motor, wheel, ping-pong ball, plastic cup, spoon, penny. The subject was instructed to put these into groups of things which were alike in some way. He was encouraged to add as many objects as possible to each group, and then questioned as to the nature of the group—why they were alike. He was told that he could use any object in a fresh group and, when he ceased producing groups spontaneously, was shown two or three objects which are commonly grouped together and asked if they were alike in any way, and whether any others would go with them. This went on until some 10-12 groups had been recorded, taking some 10-15 minutes.

On the basis of previous studies (Annett, 1959; Kagan, *et al.*, 1963) the following types of reasons were distinguished:

A.—Mere enumeration or no explanation.

B.—Associations, whether based on activity or descriptive features (e.g., match used to light cigarette).

C.—Common use by the subject, or activity of the objects; usually defined by verbs (e.g., to play with, grow on a tree, can be eaten, to do with smoking).

D.—Common concrete qualities, usually defined by nouns (made of iron, same colour, animals).

14 *Environmental Handicaps and Intellectual Development*

E.—Class names, usually defined by adjectives, or more abstract inferences (e.g., metal, vegetable, long, pointed, round, etc.).

A and B explanations were scored 0, C and D 1, E 2. For each group the score was multiplied by the number of objects it contained (1 object being deducted if the group was 'forced,' i.e., presented by the tester). The average score per group was taken i.e., the total for all groups divided by the number of groups.*

This was a task which West Indian children did relatively well, obtaining a median quotient of 90, presumably because they could describe or explain their groupings in very simple language. Previous workers, either using concrete objects (Price Williams, 1962) or Kohs Blocks (Jahoda, 1956) have likewise found that retardation of intellectual development in certain pre-literate cultures cannot be identified with Goldstein's inability to abstract.

Piaget Battery. A battery of 31 items was selected, representing a wide range of developmental concepts as described in Piaget's writings. Most of these items have been further explored by Lovell (1961), Hyde (1959) and others, and were known to show a clear progression in level of response between the ages of about 6 to 11 years. They were also chosen for brevity and for ease in evaluating the response without extensive 'clinical' enquiry. In other words, they were given very much like a Stanford-Binet scale (except that all Subjects did all items), with just sufficient questioning to ensure that the Subject had or had not grasped the concept.

The items were grouped into thirteen or fourteen sets and scored for the number of errors in each set. A brief description of the sets and their component items, together with comparative results, follows:

(a) *Time Concepts.* "What day is it today?" (unscored).

1.—"What is the day after tomorrow?" Percentage of errors or unknown: E (English) 11 per cent., WI 28 per cent.

2.—"What day is it in X?" (a well-known town 10-30 miles away). A different day was given by 22 per cent. E, 18 per cent. WI.

3.—A watch is shown with minute hand between the 5's (e.g., 10.17). "What time is it?" Greater than 3 minutes wrong: E 37 per cent., WI 38 per cent.

4.—"What time is it in X?" A different time: E 18 per cent., WI 26 per cent.

5.—"Why has a clock or watch got two hands?" Press until clear whether realises that minute hand goes faster than hour. Failure in 28 per cent. E, 56 per cent. WI.

Total error score on this set ranges from 0 to 5. Means: E 1.16, WI 1.66.

(b) *Left and Right.* "Show me your left ear."

6.—"Which is my right ear?" Wrong responses: E 28 per cent., WI 40 per cent.

7.—Lay a pencil, penny and rubber in a line. "Is the rubber to the left of the penny or the right? Is the rubber to the left of the pencil or the right?" Similarly for each object, six questions in all. Score $\frac{1}{2}$ for each error. Numbers with any error: E 39 per cent. WI 70 per cent.

Total error score 0-4. E mean 0.91, WI 1.52.

* A simpler scoring technique will be used in future based on 2 points for each genuine similarity, 1 for less effective or more concrete resemblances, 0 for associations. The total score, after multiplying by the sizes of groups, was found to correlate 0.457 in the English sample, with the Matrices test, whereas the original scoring, described above correlated 0.339.

(c) *Equidistant Counters.* One red counter in front of tester, one in front of Subject, 2 ft. apart. Scatter 5 green ones.

8.—"Take a green counter and place it so that it is as far from your red one as it is from my red one. Place another one so that it is as far . . .," and so on. "And if I gave you a large number of green ones, how far would the line of counters go?"

Score 2 for no conception of infinity or placement on a straight line, 1 for slight attempt at linear placement. Means: E 1.05, WI 1.36.

(d) *Logical Inclusion.*

Lay out four white squares, two blue squares and three blue circles in clumps as shown.

	WS		BS
WS	WS		BS
WS			
		BC	
		BC	BC

9.—"Are all the blue ones circles? Why?" Errors: E 4 per cent., WI 8 per cent.

10.—"Are all the squares white? Why?" 8 per cent. and 14 per cent.

11.—"Are there more circles or are there more blue things? Why?" 58 per cent. and 54 per cent.

12.—"Are there more blue things than there are squares, or the same, or fewer? Why?" 72 per cent. and 68 per cent.

Right answers are often given for the wrong reasons; these are scored as errors. Total errors 0-4. Means: E 1.42, WI 1.44.

(e) *Visualizing Water in Tilted Bottle.* A corked lemonade bottle, half-full of coloured liquid, and an outline drawing, half-size. The Subject draws in the level of the liquid.

13.—Bottle is tilted to 45° behind a screen, with only the top showing. "Here is a picture of what the bottle is like now. You draw where the water would be if you could see it." Lines more than 10° from the horizontal marked wrong. E 48 per cent., WI 46 per cent.

14.—Bottle is laid horizontally with only the cork visible. The question is repeated, with another drawing. E 20 per cent., WI 42 per cent.

Range of errors 0-2. Means: E 0.68, WI 0.88.

(f) *Conservation of Volume of Liquid.* Show bottle upright again and an outline drawing of a dish 3-in. wide \times 2½-in. high.

15.—"If I poured the water from the bottle into this dish, draw where it would come to." The better Subjects draw a line ½-in. to 1½-in. from the bottom. 2-2½-in. counted as wrong. E 17 per cent., WI 42 per cent.

16.—"Would there be more water in the dish than in the bottle, or less . . . , or the same . . . ? Why?" Wrong answers are about equally divided between more in the dish, or in the bottle. E 8 per cent., WI 42 per cent.

17.—Two transparent jars 3-in. diameter \times 2½-in. high are half-filled with water. Subject adjusts till equal. One glass 2-in. diameter \times 5½-in. high; one dish 4-in. diameter \times 1½-in. high.

"Are you sure we both have the same? I'm going to pour mine into this glass. Now have we got the same, or have you got more, or less than me? Why?" E 6 per cent. and WI 18 per cent. say you have more because your glass is taller.

18.—Return water to the jar. "Now I'll pour mine into this dish. Have we got the same, or have you got more, or less than me? Why?" Answers about equally divided between you have more because dish is broader, and I have more because my jar is taller. E 6 per cent., WI 18 per cent.

Question 15 (involving a picture rather than actual jars) can be scored separately or combined with Nos. 16-18, with which it correlates only moderately. Total errors 0-4. Means: E 0.37, WI 1.20.

(g) *Conservation of Amount of Plasticine.* Two balls the same colour, 1½-in. diameter. Subject adjusts till he agrees that they are equal.

19.—"You roll your's into a sausage" (up to 3-in. length). "Now have you got

more plasticine, or have I, or have we both the same amount? Why?" All wrong answers said that the sausage was more. E 5 per cent. WI 22 per cent.

20.—"Roll your's back into a ball. Have we got the same now?" Show the two jars of water. "If we dropped our balls into the water, would it go up or down, or stay the same? Why?" Errors: E 3 per cent., WI 2 per cent.

21.—"Now make your plasticine into a plate" (about 2-in. diameter). "If I put my ball into this jar, and you put your plate into your's, will the water in my jar rise more than the water in yours, or will it rise less, or will it rise the same amount? Why?" Approximately equal numbers state mine rises more because my plate is flat or broad, or yours rises more because your ball is larger or heavier. E 55 per cent., WI 44 per cent.

Total errors 0-3. Means: E 0.63, WI 0.68.

21a*.—"Can you fix your ball so that it will float? Make something with it so that it will stay on top of the water." If he fails, make a bowl about $1\frac{1}{2}$ -in. across by $1\frac{1}{2}$ -in. deep and show that it floats.

21b*.—"Why does it float now?"

(h) *Visualising Insect.* Show a circle $2\frac{1}{2}$ -in. diameter.

22.—"Here's a drawing of one of the jars on its side. I'm going to draw a little insect on top." Differentiate head and tail clearly, head pointing left. "Now he starts walking round the outside edge, like this. You draw for me what he would look like when he gets round to here." Put a dot on the bottom rim, so as not to suggest inside or outside.

Insect drawn partly or wholly inside the circle, or legs or tail downwards: E 27 per cent., WI 32 per cent. Drawn in reverse, i.e., head pointing left: E 29 per cent., WI 72 per cent. Score 0-2. Means: E 0.56, WI 1.04.

(i) *Number Concepts.* Three cards 2-in. high \times $1\frac{1}{2}$ -in., showing the numbers 2, 3, 5. Present the 2 and 3 side by side. "Tell me the biggest number you can make using these numbers." Explain why 5, 6 or 23 are wrong, and if necessary, shift the cards (not scored).

23.—Present 3 2 5. "Now what's the biggest number you could make with these?" Do not allow him to shift the cards. Errors: E 14 per cent., WI 42 per cent.

24.—Present 2 8 4 9 3 on a single card of the same height. "What's the biggest number you could make with these?" Errors: E 27 per cent., WI 68 per cent.

Total 0-2 errors. Means: E 0.47, WI 1.10.

(j) *Conservation of Length of Rods.* Two rods 6-in. \times $\frac{1}{4}$ -in. Place these parallel, longways, in front of S. "Are they the same length?"

25.—Push the rod which is nearer to the Subject $\frac{1}{2}$ -in. to the right, so that it projects beyond the other. "Tell me which one is the bigger, or are they the same? Why?" Most wrong answers say that the moved rod is the longer: E 5 per cent., WI 40 per cent.

26.—Place the same rod at 45° to the stationary one, its foot half-way along. "Now which one is the larger, or are they the same? Why?" Again the moved one is most often judged longer: E 6 per cent. WI 36 per cent.

Total 0-2 errors. Means: E 0.11, WI 0.76.

(k) *Transferring Dot to Reversed Sheet of Paper.* S is given a blank sheet of paper $5\frac{1}{2}$ -in. across \times 8-in. high. The tester has a similar one with a clear black dot 3-in. from the bottom, 2-in. from his right-hand side. A ruler and a piece of string (6-in.) are visible on the table.

27.—"You see the dot on my paper. I want you to draw one like it on your paper in just the same position. Try and draw it so that if you put your paper on mine, your dot would be just on top of mine." If he does not spontaneously do so: "You can use the string or ruler if you like."

* These questions were not actually used, but appear to be useful additions in future.

Places dot in the middle or left-hand side of his paper : E 51 per cent., WI 82 per cent.

Uses ruler or string for one dimension only, or for alignment only (i.e., positioning his dot opposite the tester's) : E 28 per cent., WI 32 per cent.

Judges purely by eye : E 44 per cent., WI 58 per cent. Score as 2 errors.

Total score 0-3. Means : E 1.67, WI 2.30.

(l) *Position and Size of Shadow.* Drawing of a street-lamp 1½-in. high and, 1½-in. to the right, a man 1-in. high.

28.—“ Here's a lamp shining at night, and here's a little man. You draw me in his shadow. Why would it be there ? ”

Score 0 : any position on the right, touching the man's feet.

Score 1 : on the right but not touching (E 28 per cent., WI 30 per cent.) ; or below the man, head or feet downwards (10 per cent. and 14 per cent.).

Score 2 : upright, on the right or above the man (10 per cent. and 22 per cent.) ; between the man and the lamp-post (13 per cent. and 16 per cent.) ; to the left of the lamp-post (2 per cent. and 6 per cent.).

29.—“ Would it be shorter than him, or longer or equal ? Why ? ” The reasons given were too complex to score. Shorter or equal answers were given by E 23 per cent., WI 30 per cent.

Total score 0-3. Means : E 1.11, WI 1.62.

(m) *Conservation of Area.* Two sheets of green blotting paper, 12-in. × 9-in., two model cows, twelve 1-in. square white blocks. “ These are two green fields which belong to a farmer. Are they the same size ? Measure off and see. The farmer has a cow in each field and they want to eat the grass. Does the cow get more to eat in this field, or in that one, or the same ? Now the farmer builds a shed in each field.” Place 1 white block in a corner of each. “ Does the cow in this field have the more to eat, or the cow in that one ? Why ? ”

30.—Place another block next to the first at the edge of one sheet, another in the middle of the second sheet. “ Does the cow in this field get more, etc. ? Why ? ” Errors : E 7 per cent., WI 18 per cent.

31.—“ Now he builds a lot of sheds.” Place remaining blocks in a row of six along one edge, scattered over the other. “ Does the cow in this field, etc. ? Why ? ” Errors : E 8 per cent., WI 22 per cent. (in about two-thirds of these the Corner-shed cow gets more).

Total score 0-2. Means : E 0.15, WI 0.40.

When all the scores were combined, the median West Indian performance fell at 86 on English norms, indicating a moderate degree of retardation. However, the results clearly varied much with different items, the greatest deficiencies occurring in :

Number Concepts.
Conservation of Water.
Conservation of Rod Lengths.
Conservation of Area.
Visualizing Insect.

The differences are negligible on :

Conservation of Plasticine.
Logical Inclusion.

And quite small on :

Time Concepts.
Equidistant Counters.
Tilted Bottle.

These differences do not seem to form any very logical pattern. However, serious backwardness in the understanding of number would be expected, whereas the sorting or abstracting ability involved in Logical Inclusion may be relatively good, as it was in the Concept Formation test. Conservation of plasticine may be good since children often play with clay and earth, although they are seriously retarded in all other types of Conservation.

Factor analysis was applied in an attempt to sort out the types of ability involved in the battery. Cuts were taken as near to the medians as possible in the scores for each set, and either coefficients of association or phi coefficients were calculated and factorized by the centroid method. On the basis of Beard's (1957) results with younger children (5 to 6-year-olds) it had been anticipated that there would be a strong general factor and little else, but specifics, i.e., that the items would mainly reflect general conceptual development. But, in each analysis two significant bipolars were extracted, their variances amounting to 48 per cent. of the total in the West Indian and 34 per cent. in the English analyses, thus indicating that the items measured at least two or three relatively distinct types of development.

In the West Indian sample, the main distinction was between predominantly numerical and verbalised concepts and the more practical concepts including conservation items. However, these results are not quoted, since the English analysis with a larger sample revealed three components which can be roughly designated as : numerical-orientational, conservation, and visualization. Rotated loadings are shown in Table 2.

TABLE 2
ROTATED FACTOR LOADINGS FOR SETS OF PIAGET ITEMS.

Item Set	Factors		
	Arithmetical	Conservation	Visualization
(i) Number Concepts63	.11	.05
(c) Equidistant Counters59	— .03	.03
(b) Left and Right57	.09	— .17
(a) Time Concepts54	.26	.08
(g) Conservation of Plasticine ..	.56	.13	— .11
(f) 15. Conservation of Liquid ..	.19	.26	.25
16-18. Conservation of Liquid ..	.16	.51	— .01
(j) Conservation of Rod Lengths ..	.50	.26	— .02
(m) Conservation of Area22	.42	.19
(d) Logical Inclusion17	.42	.22
(e) Visualizing Water in Tilted Bottle38	.07	.59
(l) Position and Size of Shadow ..	.18	— .08	.50
(h) Visualizing Insect16	— .08	.30
(k) Transferring Dot to Reversed Sheet of Paper13	.01	.30

In this table, heavy type is used to draw attention to loadings of psychological interest, rather than to statistical significance. However, with a population of 100, loadings of 0.20 upwards are likely to be statistically significant.

The first factor is still a general one, but loads the numerical items and orientation of handedness and time most strongly. The second loads all the conservation items (including time concepts), though the saturation for 'plasticine' is very small, and Logical Inclusion comes in this group rather than in the numerical group. It would appear then that West Indian boys are relatively advanced in the orientational aspects of Factor I (though retarded in number understanding). They tend to be most backward on Factor II. Conservation of plasticine seems to contradict this, but then this test has scarcely any loading on Factor II. They are also decidedly weak on three of the four components of the third factor, which involves visualising various phenomena and putting them on paper.

Further investigation of the nature of the Piaget factors was made by rescoring the English boys for these three groups of items, and scoring the West Indian responses for the two categories of items indicated above, and correlating these part-scores with the other tests. The results are described in a second article.

ACKNOWLEDGEMENTS.—The writer wishes to express his gratitude first to the Association for Aid for Crippled Children (New York), which provided the funds for this and for further researches on the same topic. The West Indian investigation could not have been carried out without the valuable assistance of the staffs of the Institute and the Department of Education in the University of the West Indies, of the University Computer Unit and of a group of students in the Education Department. In this country, sincere thanks are offered to the Education Officers, and to the Head teachers of the schools which took part.

IV.—REFERENCES.

- ANNETT, M. (1959). The classification of instances of four common class concepts by children and adults. *Brit. J. Educ. Psychol.*, 29, 223-236.
- BEARD, R. M. (1957). *An Investigation of Concept Formation Among Infant School Children*. Ph.D. Thesis, University of London.
- BERNSTEIN, B. (1961). Social structure, language and learning. *Educ. Res.*, 3, 163-176.
- DAWSON, J. L. (1963). *Psychological Effects of Social Change in a West African Community*. Ph.D. Thesis, University of Edinburgh.
- FRASER ROBERTS, J. A., and DUNSDON, M. I. (1955). A study of the performance of 2,000 children on four vocabulary tests. *Brit. J. Stat. Psychol.*, 8, 3-15.
- GOLDSTEIN, H., and SCHEERER, M. (1941). Abstract and concrete behavior. *Psychol. Monogr.*, 53, No. 239, pp. 151.
- HUDSON, W. (1961). Pictorial depth-perception in sub-cultural groups in Africa. *J. Soc. Psychol.*, 52, 183-208.
- HYDE, D. M. (1959). *An Investigation of Piaget's Theories of the Development of the Concept of Number*. Ph.D. Thesis, University of London.
- IRVINE, S. H. (1964). *Testing Abilities and Attainments in Africa*. Paper read at Brit. Psychol. Soc., Educ. Sec. meeting.
- JAHODA, G. (1956). Assessment of abstract behavior in a non-western culture. *J. Abn. Soc. Psychol.*, 53, 237-243.
- KAGAN, J., MOSS, H. A., and SIGEL, I. E. (1963). Psychological significance of styles of conceptualization. *Monogr. Soc. Res. Child Develpmt.*, 86, Vol. 28, 73-124.
- LOVELL, K. (1961). *The Growth of Basic Mathematical and Scientific Concepts in Children*. London: University of London Press.
- MACARTHUR, R. S., IRVINE, S. H., and BRIMBLE, A. R. (1964). *The Northern Rhodesian Mental Ability Survey, 1963*. Lusaka: Rhodes-Livingstone Institute.
- MCCONNELL, J. (1954). Abstract behavior among the Tepehuan. *J. Abn. Soc. Psychol.*, 49, 109-110.
- MACKAY, G. W. S., and VERNON, P. E. (1963). The measurement of learning ability. *Brit. J. Educ. Psychol.*, 33, 177-186.
- MANLEY, D. R. (1963). Mental ability in Jamaica. *Soc. and Econ. Stud.* (University of the West Indies, Institute of Social and Economic Research), 12, 51-71.

20 *Environmental Handicaps and Intellectual Development*

- MISCHEL, W. (1958). Father-absence and delay of gratification : cross-cultural comparisons. *J. Abn. Soc. Psychol.*, 56, 57-61.
- PILLINER, A. E. G., SUTHERLAND, J., and TAYLOR, E. G. (1960). Zero error in Moray House 11+ Verbal Reasoning Tests. *Brit. J. Educ. Psychol.*, 30, 53-62.
- PRICE-WILLIAMS, D. R. (1962). Abstract and concrete modes of classification in a primitive society. *Brit. J. Educ. Psychol.*, 32, 50-61.
- REID, L. H. E. (1961). *The Construction and Standardization of Attainment Tests for Secondary School Selection in Jamaica*. M.A. Thesis, University of London.
- SCHWARZ, P. A. (1961). *Aptitude Tests for Use in the Developing Nations*. Pittsburgh, Pa.: American Institute for Research.
- SEARS, R. R., MACCOBY, E. E., and LEVIN, H. (1957). *Patterns of Child Rearing*. Evanston, Ill.: Row, Peterson.
- SHAPIRO, M. B. (1954). An experimental investigation of the Block Design Rotation Effect. *Brit. J. Med. Psychol.*, 27, 84-88.
- VERNON, P. E. (1960). *Intelligence and Attainment Tests*. London : University of London Press.
- VERNON, P. E. (1961a). *The Structure of Human Abilities* (2nd edit.). London : Methuen.
- VERNON, P. E. (1961b). *Selection for Secondary Education in Jamaica*. Kingston, Jamaica : Government Printer.
- VERNON, P. E. (1962). Intellectual development in non-technological societies. *Proc. XIV Intern. Cong. Appl. Psychology*, Vol. II, pp. 94-105. Copenhagen : Munksgaard.
- VERNON, P. E. (1963). *Personality Assessment—A Critical Survey*. London : Methuen.
- WITKIN, H. A., et al. (1962). *Psychological Differentiation : Studies of Development*. New York : John Wiley.

8-409
8.9.70

COLOUR AS AN AID TO VISUAL PERCEPTION IN EARLY READING

By J. KENNETH JONES.*

SUMMARY. This paper describes the results of an investigation into the value of colour as an aid to the visual discrimination of words and letters. One-hundred-and-ten nursery school children were given a pair of matching tests, one in black and one in colour, of English reversal letters, and this was followed by another pair of matching tests in black and in colour of English words transposed into an unfamiliar script.

The aim was not only to assess the differences between the scores in the black and coloured version of each test, but to see :

- (i) Whether there was any significant sex differences in the visual perception of black and coloured material.
- (ii) Whether there was any significant difference by age.
- (iii) What was the extent of colour matching in disregard to differences of shape and orientation.
- (iv) If there was any significant difference between preferences for the black or the coloured test material.

The results showed much higher scores for the coloured version which was also preferred by the children. There were no no significant sex difficulties.

I.—INTRODUCTION.

THE first major problem facing infants who are learning to read is the visual discrimination of words and letters. Research into this problem has been restricted to traditional reading material and nothing has been done to measure any visual benefits of non-traditional reading material, such as Pitman's Initial Teaching Alphabet, or various schemes of phonetically coloured words.

The absence of such investigations is probably due to the importance attached to the phonetic rather than the visual aspects of such schemes. The present investigation was prompted by what appeared to be visual as well as phonetic benefits derived by 5 and 6-year-old children who were being taught the author's 'Colour-Story Reading'—a system based on fifty-three colour symbols, each representing a distinct sound which the children learn, incidentally, through teacher-read stories which provide concrete images of the abstract phonetic concepts.

Nursery school children were chosen for the experiment, not because the author was motivated by any desire to see how early children are ready to read in colour, but because nursery school children are sufficiently mature to undertake the visual matching tests yet not old enough to have acquired reading habits which might influence their responses.

The sample was taken from two L.E.A. nursery schools (N.100) and a private nursery school (N.10). Boys predominated at the three schools, and of the total sample of 110, seventy were boys. The ages of the total sample ranged from 3.5 to 5.1 with a mean of 4 years 3.4 months.

The girls, with a mean age of 4 years 3.8 months, were very slightly older than the boys, who had a mean age of 4 years 3.2 months.

* Now at London University Institute of Education.

II.—METHOD.

The test apparatus was devised by the author with the purpose of securing the children's interest, but without diverting attention from the tests. The basic apparatus was a tachistoscope made from a rectangular piece of white card measuring 8 inches at base and 12 inches tall. In the centre of the card, 1 inch from the base, was a rectangular cut-out, $1\frac{1}{2}$ inches long by 1 inch high. Behind this aperture was a horizontal grooved slot, along which a white card was passed which contained the test items. Above this aperture, and covering most of the card, was drawn the outline of a house front. Apart from the roof, the house was divided into six equal rectangles, in three pairs, representing three storeys, with two rooms in each storey. In each room was a cut-out rectangle representing a window, each window being the same size as the cut-out aperture at the bottom of the card. From the top of the card, two strips of white card containing the six items of test material could be slid into position behind the six windows. The children were asked to match the items which appeared one-by-one at the bottom aperture with the same items in the house.

To make sure that the children understood what they had to do, and were able to make visual matching choices, a preliminary test was given which was not scored. The six items in this test were: black square, red triangle, blue quarter moon, black circle, black cross, and green diamond.

The procedure was to show the children the empty house, and feed in the test material behind the six windows with the explanation: "These people live in the house. There is one in each room, and they are all looking out of their windows." The same six items, which were arranged in random order, were then fed through the bottom aperture on a single strip of white card. The author made no reference to shape, or colour, or black. But to help some children, particularly the younger ones, to understand what was required, the author would point to an item in the bottom aperture and say: "This is a doggie. There is one doggie in the house, and only one doggie, and it looks just like this. Can you point to the room where the doggie is? Which window has the doggie?" No child objected to a red triangle being identified as a dog, or a black cross as a mouse, etc. The children were given a pointer to indicate the appropriate windows. A paint brush was the most satisfactory, as it avoided the possibility of scratching the test material.

After exposing the six items, and correcting any wrong choices, the preliminary test material was run through again from the opposite direction. An unaided maximum score was required to avoid elimination. Five children, mostly between the ages of 3.0 and 3.6 failed. The other 110 children made the six choices correctly, and were given the scored tests.

The procedure in the scored tests was similar to that in the preliminary test, except that the author did not correct any errors, nor indicate what the scores were.

The scored test material was inserted from the top of the card so that it occupied the six windows in the house. The children were asked to notice that all the six items were different. Then a strip of white card containing the identical six items was fed horizontally across the aperture at the base of the card. The author asked: "Where does this one live in the house?" or "Which window is this one looking out of?" or "Where does he live?" With most children it was unnecessary to repeat these instructions with the appearance of each test item, as they soon realised the matching procedure that was

required. The change-over from one test to another was explained by saying : " Now some new people are coming to live in the house." The author made no verbal references to shape or colour or black or words or letters.

The first pair of scored tests contained the six English reversal letters : p, q, u, n, d, b, with one letter in each window. In one test the letters were in black and in the other they were in colour. Their colours were : p u—red ; d n—green, b q—blue, which are the same as in the author's coloured reading scheme.

The second pair of scored tests, one in black and the other in colour, comprised the six words : sat, saw, say, was, yes, yet. But they were not given in English characters. The words were transposed letter-for-letter into an unfamiliar script (a mixture of Greek and ancient Cypriot Syllabary) with each unfamiliar letter having similar characteristics to its English counterpart, letter y being represented by a below-the-line letter, t represented by an above-the-line letter, w represented by a letter comprising straight lines, and s, a, and e represented by curved letters.

In the colour version of this word test, each word had the identical colour which its English equivalent would have in the author's reading scheme. This scheme uses only four colours, red, green, blue and black, and all four colours were used in the word test. Two of the six words contained coloured backgrounds, on which the letters were presented in black. The last two letters in the equivalent of ' saw ' were written on a square blue background (' or ' sound), and the last letter in the equivalent of ' say ' was written on a round blue background (silent letters). This meant that an average of one word in three contained a coloured background, which is roughly the incidence of coloured backgrounds when English words are written in the author's colour code.

The purpose of using an unfamiliar script was the same as that for choosing nursery school children : to eliminate, as far as possible, learned reading responses.

In the black and coloured version of each pair of tests, the items appeared in a different order in the house, and were presented in a different order at the base aperture. Items with similar characteristics of colour or shape were separated as far as possible.

All 110 children were given all four tests. Roughly one-half (N.54) received them in the order : coloured letter test, black letter test, coloured word test, black word test. The other half (N.56) received them in the order : black letter test, coloured letter test, black word test, coloured word test.

After completing the four tests, each child was shown the black and coloured letters and asked " Which do you like the best, these or these ? " Then the same question was asked about the black and coloured words.

III.—RESULTS.

In general, the tests seemed to be reasonably popular, and no child appeared disinterested at any stage of the test procedure. In a few cases the tests were too popular because several children wanted to ' play the game again.' The author noticed that when children had difficulty in matching the items, they tended to point at random, usually favouring the items on the bottom storey of the house. The total test time per child, including the preliminary test, was about eight minutes.

Table 1 classifies the children according to their scores in the four tests.

TABLE 1
CHILDREN CLASSIFIED ACCORDING TO THEIR SCORES.

Test	Sex	0	1	2	3	4	5	6	Total	Mean
Black letters	Girls	1	6	8	14	9	1	1	111	2.77
	Boys	4	8	23	13	15	5	2	190	2.71
	Total	5	14	31	27	24	6	3	301	2.74
Coloured letters	Girls	0	0	1	2	7	14	16	202	5.05
	Boys	0	1	2	4	13	18	32	351	5.01
	Total	0	1	3	6	20	32	48	553	5.03
Black words	Girls	8	9	14	5	4	0	0	68	1.70
	Boys	11	24	19	11	4	1	0	116	1.66
	Total	19	33	33	16	8	1	0	184	1.67
Coloured words	Girls	1	0	6	1	8	16	8	175	4.37
	Boys	1	3	2	9	21	18	16	304	4.34
	Total	2	3	8	10	29	34	24	479	4.35

The girls tended to score slightly higher than the boys, but they were slightly older. There was no significant sex difference in the results of any of the four tests.

The differences between the scores in the black tests and in the coloured tests were highly significant. The superiority of the scores in the colour tests is greater than can be seen by comparing the respective mean scores, because so many children achieved a maximum score in colour and a random score in black.

TABLE 2
NUMBER OF CHILDREN SCORING MAXIMUM AND MINIMUM.

Test	No. of Children with maximum score	No. of children scoring 1 or nil.
Black letters	3	19
Coloured letters	48	1
Black words	0	52
Coloured words	24	5

There was little evidence of colour matching in disregard to shape and orientation. Children who matched by colour would tend to achieve a 50 per cent. score in the coloured letter test because there were two items in each colour. It is, therefore, possible to estimate the extent of colour matching by the number of errors. Green letters d and n were wrongly matched on eleven occasions compared with 195 correct matchings, and red letters p and u were wrongly matched on four occasions compared with 205 correct matchings.

This suggests that the 400 correct matchings of the letters d, n, p, u contain about fifteen (3.75 per cent.) colour matching choices, to balance the same number of incorrect colour matchings. Possibly the percentage is even less, because some of what appear to be colour matching choices may be purely random choices.

TABLE 3
RESPONSES IN COLOURED LETTER TEST.

Stimulus	Response					
	d	u	b	p	n	q
d (green)	93	—	6	1	8	2
u (red)	1	106	—	1	1	1
b (blue)	3	—	76	1	—	30
p (red)	2	3	2	99	—	4
n (green)	3	5	—	—	102	—
q (blue)	7	2	21	3	—	77

The blue letters b and q are alike not only in colour, but also in shape, and differ only in orientation. The importance of shape and colour is indicated by the relatively low scores in matching b and q.

The colour matching in the letter test might have been even smaller had not the preliminary test included only three coloured items, each of a different colour which possibly led a few children to suppose that colour was the only matching characteristic required in the subsequent tests. Only one of the 110 children, a girl aged 3.7, achieved a higher score in the two black tests than in the two colour tests, and this appeared to be due to colour matching.

The influence of age on visual matching ability is shown in Table 4 which compares the scores of the oldest children with those of the youngest children.

TABLE 4
SCORES IN THE YOUNGEST AND OLDEST GROUPS.

Test	Group	0	1	2	3	4	5	6	Total	Mean
Black letters	Young	2	2	15	5	3	3	0	74	2.47
	Old	2	4	3	6	14	1	1	95	3.06
Coloured letters	Young	0	0	3	3	4	10	10	141	4.70
	Old	0	1	0	1	3	6	20	166	5.35
Black words	Young	6	10	9	5	0	0	0	43	1.43
	Old	6	7	9	3	5	1	0	59	1.90
Coloured words	Young	1	2	3	5	6	9	4	116	3.87
	Old	0	1	1	1	8	9	11	149	4.81

The youngest group consisted of ten girls and twenty boys, with an age range of 3.5 to 4.1, and a mean of 3 years 10.4 months. The oldest group contained ten girls and twenty-one boys with an age range of 4.9 to 5.1, and a mean of 4 years 10.3 months. There was thus a difference of one year in the mean ages of the two groups.

The differences between the scores of the two groups on the tests was highly significant. Table 5 indicates the relationship between the scores of the older children in the black tests and the scores of the younger children in both tests.

TABLE 5
MEAN SCORES IN RELATION TO AGE.

	letter tests	Per cent.	word tests	Per cent.
Youngest group's mean in colour ..	4.70	100	3.87	100
Oldest group's mean in black	3.06	26	1.90	19
Youngest group's mean in black	2.47	0	1.43	0

The mean scores of the oldest children's black tests were only about a quarter and a fifth of the distance between the youngest group's mean scores in black and in colour. The relatively slow advance of the oldest group's mean scores in the black tests is even smaller than it appears, especially in the letter matching tests, where one-third of the children in the youngest group attained maximum score in colour.

Analysis of the complete results showed no significant differences in the mean scores of children at the three schools, and no significant differences in relation to the order in which the children received the four tests.

Table 6 classifies the children in accordance with the type of preference they expressed. Two boys have been omitted from this classification because, although they expressed a preference for the letters (both preferred colour), they were unable to choose between the black and coloured words.

TABLE 6
PREFERENCES IN RELATION TO SEX AND AGE.

Sex or Group	letters black	words black	letters black	words colour	letters colour	words black	letters colour	words colour
Girls	1		6		8		25	
Boys	4		10		16		38	
Total ..	5		16		24		63	
Youngest	1		4		6		18	
Oldest ..	1		4		8		17	

The difference between the black and colour preferences of the total sample was highly significant. There was no significant difference in preferences between the youngest and oldest groups, or between the boys and the girls.

IV.—RECAPITULATION AND CONCLUSIONS.

Matching the black letters and black words was at least three times as difficult as the same task in colour, even allowing for colour matching. The absence of any sex differences may seem strange in view of the general opinion

that girls learn to read earlier and more readily than boys. But this supports the findings of Goins, although her battery of visual perception tests were not so closely allied to actual reading tasks as the present investigation, nor did they involve colour, and they were given to 6-year-olds.

The oldest group achieved significantly higher scores in the tests than the youngest group. This suggests that most children gradually overcome an early tendency towards reversal errors without instruction. It supports the findings of Krise that reversals in reading is a problem in space perception.

As none of the children had any previous acquaintance with the author's Colour-Story Reading, the experiment avoided the "Hawthorne Effect." And if the results were influenced by any earlier learning of letters or words, this would favour the black tests.

The strong preferences in favour of the coloured test material is important because of the value of motivation in early reading. The extent of the preferences for colour is even more impressive because it received no support from phonetic considerations.

V.—REFERENCES.

- GOINS, J. T. (1958). *Visual Perceptual Abilities and Early Reading Progress*. Univ. of Chicago Supplementary Monographs No. 87.
JONES, J. K. (1965) Research Report *New Education* 1, 4.
KRISSE, E. M. (1949). Reversals in reading: a problem in space perception? *Elementary School Journal* XLIX.

(Manuscript received 24th June, 1964.)

THE EFFECTS OF ABSENCE ON PRIMARY SCHOOL PERFORMANCE

By J. W. B. DOUGLAS AND J. M. ROSS
(*M.R.C. Unit, London School of Economics*)

SUMMARY. Detailed absence records are available for 3,273 primary school children from the National Survey of Health and Development. They are grouped in seven categories, according to the history of school attendances between $6\frac{1}{2}$ and $10\frac{1}{2}$ years. In the upper middle classes the eleven-year mental ability and school performance test scores are unaffected by the amount of absence and by the age distribution of the absences. In all other social classes, considerable effects are recorded, children who are consistently absent or who are often absent in the last two years at primary school making low scores at eleven, and showing a relative deterioration in score between 8 and 11 years. On the whole, children who are often away in the first two years but make good attendances in the subsequent years catch up, but not if they come from the lower manual working class or go to primary schools that have a poor academic record. Even after allowing for the influence of absence on test performance, the children who are often absent get fewer grammar school places than expected; this is particularly noticeable for the manual working class children.

I.—INTRODUCTION.

AN opportunity to observe the effects of absence on primary school work is provided by the National Survey of Health and Development.* This is a continuing study of a sample of the children born in Great Britain in the first week of March, 1946. The sample comprises all the children born during this week in non-manual and agricultural workers' families, but only one-quarter of the rest of the births (illegitimate babies and twins are excluded). When these children were aged between $6\frac{1}{2}$ and $10\frac{1}{2}$ years, the primary school teachers kept running records of absences on forms specially supplied for this study. During the first of these four years, that is to say when the children were between $6\frac{1}{2}$ and $7\frac{1}{2}$, the reason for each absence was checked with the mothers by the school nurses, who visited the homes to do so. In the later years, only absences of more than one week's duration were checked in this way.

Earlier studies have suggested that the actual amount of time spent at school, within wide limits, has little effect on the standard of work done by healthy children. For example, after eight years at school, American children, who have enjoyed five months' holiday a year, were as advanced in their studies as those who had been given only four months' holiday and so spent the equivalent of an extra year at school.† In Great Britain, the independent schools have approximately a month's more holiday a year than those at the local educational authority schools, and this holds in the primary school years as well as later, yet there is no evidence to show that their pupils fare worse, indeed, one would infer the opposite from examination successes. These children however, were not subjected to the interference with learning which results

*A full account of the educational progress of these children in the primary schools and of their selection for grammar school, is to be found in *The Home and the School*, by J. W. B. Douglas, MacGibbon and Kee (1964). For a description of the background of this study, see *Children under Five*, by J. W. B. Douglas and J. M. Blomfield, Allen and Unwin (1958).

†National Society for the Study of Education: *Twenty-seventh Year Book, Part II*, Public School Publishing Co., Illinois (1928).

when a child, after being away from school during term time, has to catch up with the work of the class. It has been shown that children who were often absent were less successful in their studies than their classmates, and that a long run of illness may be a cause of backwardness by keeping children away from school and, when they return, by leaving their health impaired.* It has also been suggested that the extent to which children fall behind depends on their character and abilities. The dull ones may be seriously hampered by minor absences, and a child who, with effort, has been keeping up with a class of children of average ability, may no longer be able to do so when he rejoins the class after a period of illness. Brighter children, in contrast, may fall behind their classmates when they return, but soon catch up with special help, or perhaps even without.

The aim of this paper is to answer the following questions :

- (a) Is school work influenced by absence and, if so, to what extent ?
- (b) Does a record of frequent absences in the primary school years influence a child's performance in the 11+ selection examination ?
- (c) Are the effects of absence on test performance as marked for middle as for manual working class children ?
- (d) Are there any qualities of the school or the home that offset the effects of absence ?

II.—INFORMATION AVAILABLE.

The majority of the absences recorded by the teachers were for trivial complaints and of short duration. It is, of course, by no means certain that the explanations given by the children to their teachers, or by their mothers to the school nurses, were always the correct ones. Minor illnesses or accidents may sometimes have been used to explain absences that were, in fact, to look after brothers or sisters, or to join family expeditions. We are concerned here, however, only with the total amount of time lost and with the number of episodes of absence from school, irrespective of the cause. In fact, by far the greater part of the time lost from school (89 per cent. of the weeks lost and 87 per cent. of the episodes) is recorded as due to illness.

Complete absence records are available for 3,273 children at primary schools in England, Scotland and Wales, for whom there is also test information at 8 and 11 years. Between the ages of $6\frac{1}{2}$ and $10\frac{1}{2}$ years, a total of 65,165 episodes of absence were recorded for these children and they lost 37,832 weeks at school. On the average, each child was away from school for a total of 11.6 weeks during these four years. As would be expected, absences are more frequent in the early years than in the later ones. These records, however, start only during the second year at primary school and so miss out the first eighteen months or so when absences from all causes were probably higher. Between $6\frac{1}{2}$ and $7\frac{1}{2}$ years, when our first year's record was kept, many absences are recorded, and 4.0 weeks of school were lost on the average, whereas in the remaining three years there is a relatively constant loss varying only from 2.4 to 2.7 weeks. The time lost from school and the average number of episodes are shown in Table 1. There are comparatively small social class differences in the amount of absence and it is not until the last two years that a gradient emerges with an excess of absence in the manual working classes.

* Burt, C. : *Causes and Treatment of Backwardness*, University of London Press (1953).

TABLE 1

DISTRIBUTION OF ABSENCE BY SOCIAL CLASS.

(a) Weeks Absent.

		Average number of weeks absent				
		Age of child at beginning of school year				All 4 years
Social Class		6½	7½	8½	9½	(6½—10½)
Middle Class	Upper ..	4.1	2.6	2.4	2.1	11.1
	Lower ..	4.1	2.5	2.5	2.2	11.3
Manual Working Class	Upper ..	3.8	2.6	2.9	2.5	11.8
	Lower ..	4.0	2.7	2.9	2.6	12.2

(b) Episodes of Absence.

		Average number of episodes of absence				
		Age of child at beginning of school year				All 4 years
Social Class		6½	7½	8½	9½	(6½—10½)
Middle Class	Upper ..	5.9	4.3	4.3	4.1	18.5
	Lower ..	5.5	4.6	4.7	4.4	19.2
Manual Working Class	Upper ..	5.7	4.8	5.3	5.0	20.8
	Lower ..	5.8	5.0	5.4	5.2	21.4

The Survey children were given two test batteries, one at 8 and another at 11 years ; these included both intelligence and attainment tests. The results of the 11+ secondary selection examinations are also known.* The tests were provided for this study by the National Foundation for Educational Research in England and Wales, who also arranged for them to be scored and checked. At 8 years there was a picture intelligence test and tests of mechanical reading, vocabulary and sentence completion. At 11, the same reading and vocabulary tests were repeated, the picture intelligence test was replaced by one of mixed verbal and non-verbal intelligence, and the sentence completion test was replaced by one of arithmetic. All these tests were standardised on the inflated survey population, i.e., after allowing for the fact that only one-quarter of the children of the manual workers and the self-employed had been sampled. The final scores for each test had a mean of 50 and a standard deviation of 10. In the present study of absences the tests are not considered separately ; instead, an average of the four scores at each age is used. The fact that the children were not tested before 8 years limits this description of the influence of early absence on school performance, since we cannot measure the extent of deterioration that may have occurred before the first tests were given. A full description of the tests and their correlations is given elsewhere.†

* When the 11+ examination results are discussed, the Scottish children are omitted.

† Douglas, J. W. B. : *The Home and The School*, MacGibbon and Kee (1964) (Appendix I, by D. A. Pidgeon).

III.—THE EFFECTS OF ABSENCE ON TEST PERFORMANCE.

The number of separate episodes of absence, and the total amount of time lost from school, are used as alternative criteria for classification in Table 2 (a) and (b). In each, the children are divided into five standard deviation groups, according to the total number of episodes or weeks of absence recorded for them between $6\frac{1}{2}$ and $10\frac{1}{2}$ years.

TABLE 2

MEAN TEST PERFORMANCE AT 11 YEARS BY SOCIAL CLASS AND AMOUNT OF ABSENCE.

(a) Weeks Absent ($6\frac{1}{2}$ — $10\frac{1}{2}$ years).

Social Class		Average 11 year test scores				
		Loss of schooling between $6\frac{1}{2}$ and $10\frac{1}{2}$ years (weeks)				
		0—4 (more than 1 SD below average)	5—6 ($\frac{1}{2}$ —1 SD below average)	7—13 (within $\frac{1}{2}$ SD of average)	14—17 ($\frac{1}{2}$ —1 SD more than average)	18—93 (over 1 SD more than average)
Middle Class	Upper ..	57.13	56.91	58.12	55.49	57.44
	Lower ..	55.33	54.18	53.52	54.14	51.77
Manual Working Class	Upper ..	50.34	51.78	51.01	49.42	47.64
	Lower ..	48.89	48.69	48.45	46.35	45.74

(b) Episodes of Absence ($6\frac{1}{2}$ — $10\frac{1}{2}$ years).

		Average 11 year test scores				
		Loss of schooling between $6\frac{1}{2}$ and $10\frac{1}{2}$ years (episodes)				
		0—9 (more than 1 SD below average)	10—13 ($\frac{1}{2}$ —1 SD below average)	14—23 (within $\frac{1}{2}$ SD of average)	24—30 ($\frac{1}{2}$ —1 SD more than average)	31—88 (over 1 SD more than average)
Middle Class	Upper ..	58.50	58.52	57.53	54.66	58.27
	Lower ..	55.10	54.85	53.83	53.55	50.70
Manual Working Class	Upper ..	52.47	49.44	50.74	50.16	47.76
	Lower ..	50.08	49.08	48.06	46.99	44.93

Whether judged by weeks or episodes of absence, those most often away from school make the lowest scores in the 11 year tests. However, the upper middle class children* are unaffected by the amount of time lost from school. Their performance in the 11 year tests is not depressed even if they had many

* For a definition of the social classes, see *The Home and The School*, pages 40-44.

weeks or episodes away. In contrast, children in the three other social classes who are often away from school do consistently badly. Their poor performance is most marked when they are grouped by episodes rather than weeks of absence. As a composite assessment based on episodes and time does not give a clearer relationship between absence and performance, the number of episodes alone is used as the basis of classification in the following analysis.

So far, only the total amount of absence between $6\frac{1}{2}$ and $10\frac{1}{2}$ has been considered: it is more revealing to look at the history of absences over the years. To do this, the children with episodes of absence amounting to half a standard deviation or more above the average of their age, referred to below as 'excess' absence, were separated in each year from those less frequently away from school. The 3,273 boys and girls for whom full absence information is available were then grouped into one of the following:

- Group 1. Few episodes throughout, i.e., no excess recorded in any of the four years (1,293 children).
- Group 2. Excess in each of the first two years ($6\frac{1}{2}$ — $8\frac{1}{2}$), but not later (143 children).
- Group 3. Excess predominantly, but not solely, in the first two years (677 children).
- Group 4. Excess in each of the last two years ($8\frac{1}{2}$ — $10\frac{1}{2}$), but not earlier (80 children).
- Group 5. Excess predominantly, but not solely, in the last two years (501 children).
- Group 6. Excess in each of the four years (291 children).
- Group 7. Excess in two years, one being between $6\frac{1}{2}$ and $8\frac{1}{2}$, and the other between $8\frac{1}{2}$ and $10\frac{1}{2}$ (288 children).

Before looking further at the effects of absence on school work, a more detailed description of the children who were frequently away from school is called for. Table 3 shows that the boys, in the early years, are slightly more often away than the girls, whereas in the later years there is an excess of absences among the girls. The average difference between the sexes over the four years studied is small; the girls lose four days more from school than the boys.

The history of absence has few social class variations. Early absences (groups 2 and 3) are more common among middle class children, whereas all other types of absence are more common among manual working class children. Rather more middle class than manual working class children make consistently good attendance at school (group 1), perhaps owing to the relatively small size of middle class families, since Table 3 also shows that children with few brothers or sisters lose relatively little time from school. In the larger families, the eldest child in particular is likely to have many episodes of absence at each stage of his primary school career.

School attendance is more satisfactory among the children whose parents take an active interest in their school work, than among those who have uninterested parents. When little importance is placed on learning, the children are often away in the last two years at primary school, though not in the earlier ones. This is so whether the parents' attitudes are assessed by their contacts with the school, their views on school leaving, or the teachers' comments on the interest they have shown.

TABLE 3

SOME CHARACTERISTICS OF CHILDREN RELATED TO PATTERN OF ABSENCES.

PATTERN OF ABSENCES.								
	Absence Pattern							Total %
	Few Episodes Through-out (1)	Excess Episodes						
		In first 2 years		In last 2 years		In all years	Early and late but not in all	
		Only (2)	Predom. (3)	Only (4)	Predom. (5)	(6)	(7)	
(a) Sex.								
Boys % ..	41.4	4.6	20.6	2.1	14.5	8.3	8.5	100.0
Girls % ..	37.5	4.1	20.8	2.8	16.2	9.5	9.1	100.0
(b) Social Class								
Upper Middle%	41.7	6.2	23.2	2.3	14.7	4.6	7.3	100.0
Lower Middle%	42.3	4.3	22.5	1.8	13.5	7.3	8.3	100.0
U. Manual %	37.7	3.6	18.8	2.9	18.4	9.9	8.7	100.0
L. Manual %..	36.9	4.2	19.2	2.9	15.6	11.2	9.9	99.9
(c) Family Size								
One %	41.8	3.7	20.1	2.6	15.1	8.1	8.5	99.9
Two %	41.2	5.0	21.6	2.0	14.7	6.8	8.7	100.0
Three %	40.8	5.1	20.5	1.7	14.7	9.5	7.7	100.0
Four or more %	35.2	3.3	19.8	3.6	16.9	11.2	10.0	100.0
(d) Parents Interest								
High %	47.1	5.2	17.4	1.6	14.0	6.4	8.2	99.9
Average % ..	38.6	5.9	25.2	2.1	13.2	6.0	9.0	100.0
Low %	38.0	3.7	19.9	2.7	16.3	10.7	8.7	100.0
(e) Academic Record of School								
Good %	43.9	4.0	21.3	1.4	14.7	7.0	7.6	99.9
Fair %	39.3	4.6	20.0	2.5	15.7	9.0	9.0	100.1
Poor %	35.1	4.4	20.8	3.3	14.9	11.3	10.2	100.0

The teachers describe the pupils who are consistently absent (group 6) as being poor workers or lazy. 23.4 per cent. of the lazy children had been consistently absent as compared with only 2.6 per cent. of the very hard workers. The other types of absence history bear little relation to the teachers' assessment of their pupils. An excess of early absences is rather more common among the hard-working children and this may reflect the effort made by some children who have fallen behind in the early years to catch up. A pattern of late or variable absences is also slightly more common among the poor workers and lazy children.

TABLE 3—continued.

SOME CHARACTERISTICS OF CHILDREN RELATED TO PATTERN OF ABSENCES.

	Absence Pattern							Total %
	Few Episodes Through-out (1)	Excess Episodes						
		In first 2 years		In last 2 years		In all years	Early and late	
		Only (2)	Predom. (3)	Only (4)	Predom. (5)	(6)	(7)	
(f) Child's Attitude to Work								
Very hard % ..	48.7	5.6	22.8	2.0	11.4	2.6	6.9	100.0
Hard %	44.3	4.6	21.6	1.8	11.4	6.5	9.7	99.9
Average %	36.9	4.0	20.7	2.7	17.9	9.6	8.2	100.0
Poor %	29.3	4.4	15.1	3.2	21.0	16.3	10.7	100.0
Lazy %	30.8	2.8	17.8	1.9	14.0	23.4	9.3	100.0
(g) Area of Primary School								
Urban	40.6	4.3	20.5	2.4	14.8	8.9	8.4	99.9
Rural	35.7	4.6	21.6	2.6	17.3	8.1	10.2	100.1
(h) Region								
Scotland % ..	46.8	3.0	14.5	2.2	17.5	9.1	6.9	100.0
North %	51.3	2.6	16.1	2.2	15.0	5.5	7.3	100.0
N.W. and Yorks %	41.3	3.5	20.3	2.8	13.7	11.2	7.2	100.0
Midlands and N.Midlands %	42.5	4.7	19.5	3.1	15.3	7.1	7.7	99.9
Wales %	29.7	2.9	23.4	3.4	15.4	13.1	12.0	99.9
East %	33.3	3.7	25.0	2.1	15.8	8.3	11.7	99.9
S. and S.W. %	34.7	6.6	26.9	2.6	13.9	6.4	9.7	100.1
London and S.E. %	33.8	5.7	21.3	1.6	16.5	10.5	10.6	100.0

The history of absence is also related to the type of school a child attends. In country schools, perhaps because of difficulties with transport, absences are rather more frequent than in the towns. In all areas the pupils at primary schools which have a good past record of gaining grammar school places have themselves a good record of attendance; fewer are consistently away (group 6) and fewer show a variable history of absence (group 7). It might be thought that the time lost from school would, for climatic reasons, increase as one moves North. This is not so; children in the Midlands, the South and Wales are more likely to be absent in the early years (group 2) than children in Scotland and the Northern part of England, and they are more likely to have a variable history of absence (group 7).

TABLE 4
MEAN TEST SCORES RELATED TO PATTERN OF ABSENCES.
MEAN TEST SCORE

Test Score	Absence Pattern						
	Few Episodes Through-out (1)	Excess Episodes					
		In first 2 years		In last 2 years		In all years (6)	Early and late (7)
		Only (2)	Predom. (3)	Only (4)	Predom. (5)		
At 8 years	51.76	50.68	50.29	50.57	49.58	47.10	50.20
At 11 years	51.85	52.38	50.30	48.78	49.31	47.05	49.49
Change in Score.	+0.09	+1.70	+0.01	-1.79	-0.27	-0.05	-0.71
Change for those with 8 yr. scores with $\frac{1}{2}$ SD of mean*	+0.09 (469)	+2.52 (47)	+0.61 (232)	-4.19 (29)	-0.34 (174)	-0.37 (109)	-0.69 (93)

* Figures in brackets show the number of children in the group within $\frac{1}{2}$ SD of mean—these are the actual numbers and not population estimates.

Test performance, as one would expect, varies with the history of absence. Table 4 gives the average test scores at 8 and 11 years of children in the seven absence groups. The highest scores at 11 are made by those who were consistently absent between $6\frac{1}{2}$ and $8\frac{1}{2}$ and had a good attendance record thereafter (group 2) and the lowest by those who were consistently away (group 6). Those in group 4, i.e., with excessive absences between $8\frac{1}{2}$ and $10\frac{1}{2}$ years only, also make low scores at 11.

More revealing are the changes in score between 8 and 11 years (line three of Table 4). The children with an excess of early absences (group 2) improve their scores, whereas those absent in the later years (group 4) fall behind. It was expected that the consistently absent children (group 6) would also deteriorate in performance and the fact that this expectation is not fulfilled suggests that, for some at least, active efforts are being made during the later years at primary school to compensate for lost class work. In contrast, a variable history of absence (group 7) is associated with a deterioration in measured ability, less marked though, than in group 4 where absence is concentrated in the last two school years.

The average test scores in Table 4 differ significantly from each other ($P < 0.01$) and the changes in test score also differ significantly ($0.05 > P > 0.01$). These changes in test score are difficult to interpret because the children in these seven absence groups start with different average scores. Those with few absences make higher scores at 8 than those with many and so, for the following reasons, the crude differences between 8 and 11 year scores under-estimate the influence of absence on test performance. When the correlation between tests is less than perfect (in the present study the co-efficient r is +0.82) those that make high scores on one test are likely to make lower scores on another, whereas those that make low scores are likely to improve. A clearer account

of the progress of children with different histories of absence is obtained if only those children (1,622*) with scores near 50 at 8 years are considered.

Line four of Table 4 has the same absence groupings but it only refers to those whose 8 year test scores fell from 46 to 54 inclusive ($\frac{1}{2}$ S.D. on either side of the mean).† When this is done, four highly significant differences ($P < 0.01$) appear. The children who were often away in the early years (group 2) improve their position subsequently, whereas those who were away in the last two years (group 4) show relatively a large deterioration. The consistently absent children (group 6) still show relatively less deterioration than those who are only absent in the later years (group 4), though they now fall 1.3 points behind those with a consistently good attendance record (group 1).

TABLE 5

MEAN TEST SCORES AT 11 YEARS BY SOCIAL CLASS AND PATTERN OF ABSENCES.

Average 11 year test scores.

Social Class		Absence Pattern						
		Few Episodes Through-out (1)	Excess Episodes					
			In first 2 years		In last 2 years		In all years (6)	Early and late (7)
			Only	Predom.	Only	Predom.		
			(2)	(3)	(4)	(5)		
Middle	Upper	57.89	58.61	56.17	55.89	58.71	57.56	57.26
	Lower	55.27	53.51	53.34	54.09	52.00	52.34	52.99
Manual Working	Upper	51.36	53.06	50.25	50.98	49.05	47.27	49.30
	Lower	49.36	50.23	47.73	44.82	47.10	44.80	47.21

At the age of 11, the upper middle class children in the complete absence information sample (see Table 5) are little affected by either the amount or the timing of their school absences. In each of the other three social classes, however, the effects of absence, as shown by the 11 year scores, are considerable. The most heavily affected children are those whose school work has been consistently interrupted. The least affected are those who were absent in the early years, but not later. The maximum social class difference in measured ability which is 8.53 points for the children with few absences, rises to 12.76 points for the children who were consistently away.

* It should be remembered that the tests were standardised on the inflated population. When this is allowed for this group forms 40 per cent. of the total of children in the original population for whom full absence data is available.

† When this restricted group of children, with measured ability from 46-54, inclusive is taken, the mean score at 8 years is closely similar for each absence group being:

Group 1 = 50.12	Group 2 = 49.90	Group 3 = 49.78
Group 4 = 50.31	Group 5 = 49.96	Group 6 = 49.70
Group 7 = 50.30		

These means are not significantly different from each other, and the changes in score between 8 and 11 years shown in the last line of Table 4, and equally the changes shown in Tables 6, 8 and 9, cannot be explained in terms of regression.

TABLE 6

CHANGE IN TEST SCORE BETWEEN EIGHT AND ELEVEN YEARS BY SOCIAL CLASS AND PATTERN OF ABSENCES.

Change in Score 8—11 years*

Social Class		Absence Pattern						
		Few Episodes Through-out (1)	Excess Episodes					
			In first 2 years		In last 2 years		In all years (6)	Early and late (7)
			Only	Predom.	Only	Predom.		
			(2)	(3)	(4)	(5)		
Middle	Upper	+3.26	+6.89	+2.40	+2.75	-0.50	+1.29	-0.43
	Lower	+2.37	+5.10	+2.25	-1.00	+0.51	+2.68	+0.76
Manual Working	Upper	+0.42	+2.86	-0.26	-5.27	+0.32	-0.67	-1.93
	Lower	+0.20	+0.17	0.00	-4.33	-0.95	-1.22	-0.69

* For those with 8 year scores within $\frac{1}{2}$ SD of mean + = improvement ; - = deterioration.

The changes in test score between 8 and 11 years (taking again only 1,622 children who at 8 made scores within half a standard deviation of the mean) are shown in Table 6. In the lower manual working class all children with excessive spells of absence deteriorate in test scores relative to those with few absences (group 1), those absent during the last two primary school years being the most affected. In the other three social classes those with early absences (group 2) improve their scores in relation to group 1, and this is specially marked in the two middle classes. The children with a variable history of absence (group 7) show a relative deterioration in score in each social class as also do those who have an excess of absences between $8\frac{1}{2}$ and $10\frac{1}{2}$ (group 4). Those consistently absent (group 6) show a relative deterioration in the upper middle and manual working classes and a slight improvement in the lower middle class.

So far we have looked at test scores. There is a further question: are children who are absent from school at a disadvantage in the secondary selection examinations? In Table 7, the proportion given grammar school places is shown for children in five absence groups; owing to small numbers, the two groups of early absences have been combined as have also the two groups of later absences.

In the middle classes, those with either a consistent or a variable history of poor school attendance (groups 6 and 7) get significantly fewer grammar school places than expected from their 11-year-old test performances. The less able children (T scores 52—60) are the most disadvantaged; the cleverer ones do not have significantly less chance of a grammar school place, even when they have been consistently absent. In the manual working classes, all of the less able children with a history of excess absence are handicapped in the 11+ selection process, and even among the cleverer ones those with an excess of absences in all years are at a disadvantage. In each class, and at each level of measured ability, those with a consistent history of absence (group 6) are at the greatest disadvantage.

TABLE 7

PERCENTAGES GAINING GRAMMAR SCHOOL PLACES IN ENGLAND AND WALES BY SOCIAL CLASS AND PATTERN OF ABSENCES.

	Ability	Absence Pattern						
		Few Episodes Through-out	Excess Episodes				Significance	
			In first 2 years	In last 2 years	In all years	Early and Late		
		(1)	(2) & (3)	(4) & (5)	(6)	(7)	χ^2	n p
Middle Class	52—60	35.0	45.4	33.0	21.1	27.1	16.50	4 <0.01
	60 & over	98.8	87.8	87.2	82.8	87.9	0.95	4 0.95 — 0.90
Manual Working Class	52—60	24.9	16.0	20.9	12.0	18.6	14.76	4 <0.01
	60 & over	86.6	75.0	90.0	68.8	78.3	17.62	4 <0.01

TABLE 8

CHANGE IN TEST SCORE BETWEEN EIGHT AND ELEVEN YEARS BY LEVEL OF PARENTS' INTEREST AND PATTERN OF ABSENCES.

Change in Score 8—11 years.*

Absence Pattern							
Level of Parents' Interest	Few Episodes Through-out	Excess Episodes					
		In first 2 years		In last 2 years		In all years	Early and late
		Only (2)	Predom. (3)	Only (4)	Predom. (5)		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	
High	+3.61	+6.00	+1.02	†	+1.30	+1.65	+0.72
Average	+1.72	+1.09	+1.11	†	-0.05	+1.00	-0.17
Low	+0.42	+2.42	+0.36	-3.33	-0.60	-0.73	-1.08

* For those with 8 year scores within $\frac{1}{2}$ SD of mean, <—> improvement.

† <10 children.

—=deterioration.

It has been shown elsewhere* that children improve their test performance if their parents take an interest in their school work and encourage them. In these circumstances it is to be expected that the effects of absence would be

* *The Home and The School.*

minimised, but Table 8 (which refers to the 1,622 children scoring within $\frac{1}{2}$ S.D. of the mean) shows that there is no evidence for this in our data. The children whose parents take much interest in their work make higher scores than the rest at both 8 and 11 years, but the effects of absence are equally evident whether the parents are interested or not. The children of parents who show a high level of interest improve rather more than the rest if they have a history of early absence (group 2); on the other hand, they show a relatively greater deterioration in performance if their history is of consistent or variable absence (groups 6 and 7).

The influence of the primary school on the test performance of children with a history of early absence, shown in Table 9, is considerable. When the school has a good academic record,* those in group 2 improve their scores by 4.5 points as compared with an improvement of only 1.5 points among children at the same type of school who have seldom been away. The middle class children, in these circumstances, make a more dramatic improvement in their scores than the manual working class. In contrast, the group 2 children who are at primary schools with a poor academic record not only fail to catch up, but fall further behind. One would like to know the extent to which the teachers and parents of children at primary schools with a good academic record make special efforts to help those that have been away; but unfortunately our information is insufficiently detailed. This is the type of important question that cannot be answered in a large-scale study of this type; it needs a more intimate approach which would be well worth pursuing.

* See page 104 of *The Home and The School* for a definition of the grouping of the primary schools.

TABLE 9

CHANGE IN TEST SCORE BETWEEN EIGHT AND ELEVEN YEARS BY ACADEMIC RECORD OF PRIMARY SCHOOL AND PATTERN OF ABSENCES.

Change in Score 8—11 years.*

Academic record of Primary School	Absence Pattern						
	Few Episodes Throughout	Excess Episodes					
		In first 2 years		In last 2 years		In all years	Early and late
		Only (2)	Predom. (3)	Only (4)	Predom. (5)	(6)	(7)
Good	+1.50	+4.46	+1.92	†	+1.40	+0.89	-0.03
Fair	+1.10	+2.78	+0.46	-3.54	-1.26	-0.92	-0.58
Poor	-0.09	-0.38	-0.49	-4.40	+0.07	-0.46	-1.45

* For those with 8 year scores within $\frac{1}{2}$ SD of mean, + = improvement.
-- = deterioration.

† <10 children.

IV.—CONCLUSIONS.

This study of the relation between the history of school absence and test performance in a national sample of 3,273 children, emphasises once more the importance of the home and the school in primary education. Absence from school impairs a child's measured ability and frequent episodes of absence are more harmful than an occasional long one. But this is not an invariable association. In the upper middle class, even children who have been consistently absent maintain their level of performance. In the best primary schools the children who have lost much time at the beginning of their school careers catch up with and surpass the rest. Children with a history of many absences are handicapped both through failing to keep up with their class mates and, at the point of secondary selection, through failing to get the grammar school places that would be expected from their performance in the tests. From this it would seem, first, that special studies are needed to see how the child who loses time from school can be helped to recover lost ground and, second, that in secondary selection account should be taken of each child's absence history and, for those who have lost much time from school, a greater reliance should be placed on measured ability and less on the past school record and the teachers' assessments.

ACKNOWLEDGEMENTS.—The National Survey of Health and Development was sponsored by a joint committee of the Institute of Child Health, the Society of Medical Officers of Health and the Population Investigation Committee. It is now being directed from the Medical Research Council Unit at the London School of Economics. In its early years the inquiry was financed by the Nuffield Foundation, and grants for continuing it through the school years have come from many sources, including the Ford Foundation and the Population Council, Inc.

We have received much assistance from the National Foundation for Educational Research in England and Wales—they supplied us with all the school tests and both checked the markings and standardised the scores. We wish to thank them for all their help, and also the Scottish Council for Research in Education for advice we have had on many occasions.

A National Longitudinal Study such as this is only possible because of the generous co-operation we have had from many people and agencies: here we would especially like to thank the directors of education, teachers and health visitors in all parts of Great Britain, who have given us information about the boys and girls included in this Survey. We should also like to thank the mothers who willingly answered numerous detailed questions, and the boys and girls themselves.

THE EVALUATION OF CERTAIN SCIENCE PRACTICAL TESTS AT THE SECONDARY SCHOOL LEVEL

BY A. A. ABOUSEIF AND D. M. LEE
(*University of London Institute of Education*)

SUMMARY. Science teachers have expressed dissatisfaction regarding both the validity and the predictive value of the conventional practical tests in science. The proposed new secondary school examinations have necessitated a rethinking of the purpose of such tests, and the investigation here described is a contribution to this field. It has been carried out as far as possible within the kind of conditions which may obtain for these new examinations.

Three forms of practical science tests were based on six hypothetical components of practical ability in science. Tests were constructed in biology, chemistry and physics. Using 196 secondary school pupils, the results from these tests were compared with scores on tests of attainment and ability in science, and with corresponding teachers' estimates. The analysis of these scores showed wide differences between schools and between identical test patterns in the three science subjects. The differences between schools were no doubt temporary facets of a new and disturbing situation. Greater theoretical interest attaches to the differences between the test adaptability of the separate subjects, and it would seem that identical test procedures cannot be used in each field.

I.—INTRODUCTION.

THIS research is part of a growing body of work which explores the uses and advantages of psychometric methods to evaluate practical educational problems. This general field differs from the more customary type of research in that the total concept and investigation of the problem is dependent on persons and factors external to the research field, and the entire study is thus set within a predetermined framework. The limitations consequent on this are part of the investigation so long as the results are intended to hold in the practical field, and to influence corresponding decisions and arrangements. The normal workaday setting cannot be replaced by the ideal experimental environment. The techniques to be employed must be the best possible in terms of this setting and for the furnishing of the kind of results required. Furthermore, the co-operation of trained teachers is of paramount importance where the outcome of the work is to be translated into action in the broad educational field. Only in this way can any research maintain such ties with the general field of Education as will be sufficiently close to maintain the maximum possibilities for a wide and competent usage of its results. It was against this general background that the present research came into being. Help was sought by a panel of Science teachers framing science examinations of the new secondary school type (the type now commonly known as the Certificate of Secondary Education, or C.S.E.). The particular problem at issue was the form of practical test to be used. The requirements were that the form chosen should be sufficiently straightforward to be administered and controlled by teachers, similar for each of the separate fields of physics, chemistry and biology, and suitable to the full examination of which it was to be a part. From the research angle, the requirements were that the tests should be mutually independent and valid for their purpose, and should bear adequate relationships to the remainder of the science assessments.

II.—OUTLINE OF PREVIOUS RESEARCH.

Peel¹ has pointed out the dearth of research evidence in the field of Science teaching, and in particular, the lack of proven data in regard to practical examinations. Ward² and Lomas³ have both been concerned with assessing the contribution of the present form of practical examination at A Level in both biology and physics. These writers were not concerned with possible reconstructions of these examinations at either Advanced or lower levels. Moreover, they were concerned with reliability rather than validity criteria.

For evidence of underlying components of practical ability in any branch of science one has to look to the philosophical writers or to teachers. Kerr⁴ has summarised the available evidence but it is somewhat incomplete. It is still necessary to analyse the functions of practical work in science in terms of the likely psychological abilities brought into play at each stage. In the conventional practical examination at any level, each pupil usually performs one or two experiments. Such a test implies at least three presuppositions. Firstly, the work involved is of one psychological structure only and depends to a large extent on the ability to give a written report on a logically complete performance. Secondly, the mark thus awarded on one tiny fraction of a large syllabus is accepted as a valid and reliable test of the student's performance. Thirdly, the pupil is assumed to be capable of persistent effort of a fairly lengthy kind.

A review of this scant evidence merely stresses the problems inherent in the present study since it offers but little guidance in relation to the main objects to be pursued.

III.—PLAN OF THE INVESTIGATION.

(i) *Subjects and materials.*

The nature of this investigation involved two kinds of subjects. On the one hand, a certain degree of pilot work was necessary in order to gain opinions and judgments on the nature of practical ability, as well as on the nature of practical work in science and related measurement techniques. For these purposes, groups of student or practising teachers were used. On the other hand, the carrying out of the investigation necessitated a choice of schools within the area covered by the examination. Having regard to the newness of the latter, and to the known limitations of science teaching at this level, schools were chosen on a volunteer basis, and provided a varied group typical of both teachers and pupils in the area represented. Two schools were used for pilot work, and seven were used in each science subject in the main investigation. The corresponding science teachers were trained as assessors. In the main investigation there were seventy two pupils in physics, sixty one in chemistry, and sixty three in biology. This made an overall total of one-hundred-and-twenty-one boys and seventy-five girls. That not all schools could offer all three science subjects was part of the state of science teaching at that time, and it was necessary to confine the main investigation to those who had studied a common syllabus in the subject concerned for at least two years. In addition, the whole sample was drawn from fifth-form pupils staying on for an extended course not connected with G.C.E., and with a probable intelligence quotient range of 100–115.

The materials used for this investigation comprised tests and questionnaires of various kinds as well as instruction sheets and assessment forms for use in the actual practical tests. Prior to the construction of the main experiment, a questionnaire on the aims and principles of science teaching, modes of assessment of practical work, and definitions of scientific ability was completed by two groups of practising and student science teachers. A further questionnaire

was prepared for use with the schools chosen for the experiment. This asked for information on science courses taught in the school, opportunities for practical work, range of available equipment and ability ranges of the candidate concerned. This information was sought in order to assess the comparability of the schools concerned.

Past researches, preliminary reading, and evidence from the questionnaire to students and teachers suggested the following six components underlying success in practical work in science, and therefore, necessitating assessment in connection with the practical tests :

- (a) The ability to identify various kinds, sizes, types and grades of apparatus and materials (name, N).
- (b) A knowledge of specific uses of materials, objects and apparatus and the ability to select those which are appropriate for specific purposes (use, U).
- (c) The ability to perform simple operations or specific steps in complex operations (performance, P).
- (d) The ability to convey to others a clear statement of observed facts (report, R).
- (e) The ability to apply knowledge to practical methods and procedures to new experiments (solving a problem, S).
- (f) A capacity for methodical and tidy working, e.g., in setting up and using apparatus (observe, O).

At the same time, discussion with teachers and others suggested a trial of three forms of practical test :

- (g) The simple tasks approach, comprising a suitable number of short questions (Type I).
- (h) The 'wholeness' approach, asking the completion of one or more experiments of the customary types (Type II).
- (i) A combination of the above two approaches (Type III).

Practical tests of all three types were constructed for each of the three science subjects, having due regard to the limitations of school conditions and equipment. Each test of Type I contained eighteen questions which were answered by the pupils on a special sheet prepared for the purpose. Each test of Type II was so arranged that one of four experiments was allocated (without choice) to one or more pupils. Again, a specially prepared answer sheet was used. Each test of Type III consisted of two parts. The first was of Type I but contained nine questions; the second was of Type II but asked about one half of the requirements. The time allowed for each test was forty-five minutes. The assessment was in each case based on the six components underlying practical work in science and outlined above. As the success or otherwise of these tests ultimately depended on the close collaboration of the science teachers concerned, careful instruction sheets were written and demonstration examinations were arranged.

Certain criterion were needed for assessing the validity of the practical examinations. The best available criteria appeared to be an attainment test (A.T.), a scientific ability test (S.A.T.), and teachers' estimates (T.E.) based on 'On-the-job performance.' The panel of science teachers responsible for the form of science examination had already agreed the following structure :

- PART A. A short-question type attainment test covering all three subjects and comprising multiple-choice and open-ended questions.
- PART B. A practical test (the object of this investigation).
- PART C. An essay-type paper.

The attainment test used as Part A was deemed suitable for use in assessing the validity of the proposed practical tests. It contained seventy-five questions covering a considerable amount of ground, and carried a time-limit of thirty-five minutes. Being a compulsory part of the full examination, such a test was already available in a form suited to the sample concerned. No comparable test of scientific ability was discovered, and accordingly one was prepared from information gained in the preliminary teachers' questionnaires. It was desired to test scientific thinking and method without recourse to the measurement of attainment, but having regard to possible related abilities from the more general field. Using scientific material, a test was constructed in four parts, and aimed to test the six ability classifications of conceptual redefinition, education of structural patterns, figural identification, judgment, length estimation and visualization. The final form of this test carried a working-time of forty minutes. For a third criterion for validation, teachers were asked to give estimates of their pupils by completing a prepared rating list. The estimates were requested on the usual five-point scale and covered the six components of practical ability outlined above.

Full details of all material compiled for this investigation are available in reference form.⁵

(ii) *Experimental Procedures.*

From the initial stages of the enquiry, all work proceeded in conjunction with the panel of science teachers responsible for the final form of examination, and all decisions were the outcome of joint discussion based on available evidence. By this method, the best of the experimental evidence could be considered in the light of the practical examination requirements.

Two kinds of pilot investigation were carried out in the Spring Term, 1962. Two groups of forty practising or student science teachers answered the questionnaires which gave the basic evidence for the construction of the scientific ability test and the practical tests themselves. Later in the same term, a preliminary version of the scientific ability test was tried out with some 150 subjects, and a final version was constructed. Meanwhile, the experimental forms for the practical tests were decided. Corresponding sets of instructions for assessors and candidates were drawn up and validated by discussion with the science teachers' panel. Answer sheets were compiled, and discussed in a similar way. Early in the Summer term a full try-out of the three practical tests took place in two schools covering all three science subjects. Attempt was made to keep as far as possible to examination conditions, at the same time checking on all details of the procedure itself, including answer sheets, instruction forms, problems of laying out laboratory equipment, and the mechanics of conducting the tests themselves. After thorough discussion of this pilot study, a public demonstration of the laboratory layouts, instructions, and mark sheets for all three types of practical tests was open to any London teacher who cared to attend. In this way, about 200 teachers were familiarised with the proposed practical examination alternatives.

After this pilot work was completed, volunteers were sought for the main investigation. All volunteer schools were used, and all answered the questionnaire which sought details of syllabus followed, laboratory equipment and capacity, and other background information. Final versions of the instructions and mark sheets to be used were sent out to the schools, and a week was arranged for the main experimental work. First, teachers were asked to give their estimates on the sheet provided. Two days were then allowed for laying out the laboratories in readiness for the examination procedure. During this time every

effort was made to ensure that each school understood the correct procedure to be followed, and that each science teacher fully understood the forms of assessment required. The three types of practical test were then carried out for each subject. In each school, the administration of the attainment test and the scientific ability tests followed later, and was carried out by the writers under normal test conditions.

IV.—STATEMENT OF RESULTS.

The initial questionnaire to both student and practising science teachers showed close agreement on seventeen items. These results provided a basis for selecting components underlying both scientific ability and practical work. The questionnaire to heads of science departments in the schools sample showed a common syllabus for at least two previous years, a considerable emphasis on practical work, and a wide range of available scientific equipment. It also indicated wide differences of both mental ability and social background amongst the pupils. These differences would be characteristic of pupils taking the new examination.

As the main investigation was essentially exploratory in nature, broad results concerning the validity of both structure and purpose were sought by analysis of variance and correlation techniques.

TABLE 1

SCHOOL DIFFERENCES IN EACH SUBJECT FOR EACH TYPE OF TEST AND EACH COMPONENT OF PRACTICAL WORK.

Subject	Type of Test	F ratio	Component	F ratio
Physics	Type I	7.96*	N	2.24
	Type II	3.37*	U	29.20*
	Type III	1.44	P	16.03*
	A.T. (P)	2.28*	S	3.53*
	S.A.T.	1.68	R	16.63*
			O	4.75*
Chemistry	Type I	12.03*	N	17.86*
	Type II	4.29*	U	3.68*
	Type III	10.67*	P	1.38
	A.T. (C)	4.05*	S	14.84*
	S.A.T.	2.72*	R	13.88*
			O	10.21*
Biology	Type I	20.71*	N	12.11*
	Type II	6.80*	U	14.79*
	Type III	4.88*	P	4.33*
	A.T. (B)	4.27*	S	8.96*
	S.A.T.	1.01	R	6.92*
			O	13.51*
All subjects combined	A.T. (general science)	6.92*	* Denotes an F ratio significant at 5 per cent. level.	
	S.A.T.	7.68*		

It will be seen that the schools involved in each subject-group differed significantly on most of the practical tests, on the A.T. (subject), and also on the S.A.T. (chemistry group) at the 5 per cent. level of significance. Table 1 also indicates that differences on the Type 1 tests are generally of higher levels

of significance than differences on the other two types. In physics and biology differences are greatest on Type I tests, with Type II tests showing the second largest differences, while in chemistry the Type III test shows the second largest difference. Furthermore, the table shows, with high confidence, that the nine schools have not been drawn from one or duplicate populations in view of the A.T. (general science) and the S.A.T. The major reason for the high level of significance connected with these two tests may arise from the actual combinations of groups studying different science subjects. On the other hand, the concept of practical work, which represents the main part of this investigation, seems to have been assimilated and developed in different ways by the schools involved, and this in itself could have produced some of these significant school differences.

It will be noted that the schools of each subject group differed significantly at the 5 per cent. level on the components underlying the assessment of practical work and the S.A.T., except in the case of the 'name' component of physics and the 'perform' component in chemistry. The overall pattern of these results suggests that each component differs in importance in the assessment of each subject. This might be a matter of differences in the organisation and uses of practical work within each school, or it might ultimately prove to be due to real differences in the abilities demanded in the different subjects.

TABLE 2

SCHOOL DIFFERENCES ON THE THREE TYPES OF PRACTICAL TESTS TAKEN JOINTLY.

Subject	F ratio	
Physics86	* As in Table 1.
Chemistry	3.17*	
Biology	7.10*	

It will be noted from Table 2 that, when the practical tests are taken together as a single variable, there are no school differences in physics, a difference significant at the 5 per cent. level in chemistry, and a 1 per cent. level significance in biology. Application of the t-test following these analyses of variance, showed that, in chemistry, the results on Type I are significantly higher than those on Type II, and in biology the results on Type III are higher than on Type I. Similarly, the A.T. (subject) and the S.A.T. differed significantly from each other in the three subject groups. Thus, it would seem that the attainment and ability tests diverge significantly in assessing success in the separate sciences, whether it be present accomplishment or the prediction of future success.

As a result of these significant differences between the nine schools, as shown in Tables 1 and 2, the sample was divided for correlational treatment under three groupings:

- (1) Each school separately.
- (2) School A for physics, school F for chemistry, school G for biology, since these were judged to be the most representative in each subject group.
- (3) A group of three schools in each subject which showed non significant differences amongst themselves, and which therefore, could be combined for further analysis of the scores.

This break-down of the sample is a product of the known unevenness of background in the schools used. It was to be expected and could not be avoided. It will be noted, however, that break-down under sections (1) and (2) above, led to very small numbers in the related sample groups. The break-down under (3) gave larger and more reliable groups. The correlation coefficients between the three types of practical test in each subject and the three criterion measures were calculated for each of the sample groups described above, and the results for the combined groups under (3) are shown in Table 3. Though the numbers in the other groups were too small to be reliable, the general pattern was similar to that shown in this table.

TABLE 3

INTER-RELATIONSHIPS BETWEEN THE PRACTICAL TESTS AND CRITERIA FOR THE THREE-SCHOOL SAMPLE.

Test type	Criterion								
	Physics group			Chemistry group			Biology group		
	A.T.	S.A.T.	T.E.	A.T.	S.A.T.	T.E.	A.T.	S.A.T.	T.E.
I	.30	.23	.57*	.55*	.17	.46*	.42	.44	.62*
II	.47*	.24	.36	-.34	.52*	.10	-.16	-.07	.02
III	.54*	.67*	.58*	.20	.26	.55*	.15	.30	.15

* Significant correlations at the 5 per cent. level.

Having regard to the main purpose of the investigation, it seemed reasonable that the type of practical test showing the highest set of relationships with T.E., and fairly high relationships with A.T. and S.A.T., might be considered as theoretically the most promising one for the assessment of practical work. From this standpoint, the evidence shown in Table 3 suggests that the Type I test is the most appropriate for physics, Type III for chemistry, and Type I for biology. Considering the results for physics only, the Type III test shows as high a correlation with the written attainment test (A.T.) as with the teacher's estimates of achievement in practical work (T.E.). A similar pattern shows with the Type II test. Thus, both these forms of test overlap with the written assessment of what has been learned, and hence they are less good forms of assessing the practical work.

Considering the results for chemistry separately, the Type III test best fits the criterion set out above. The Type I test again overlaps with the written attainment test and is, therefore, not so suitable. The Type II test appears to show significant overlap with the scientific ability test (S.A.T.) only, but this is not sufficient in view of its low correlation with the teachers' estimates. On the other hand, consideration of the biology results shows clearly that the Type I test is the most appropriate form of practical assessment. Clearly, since the numbers involved in the three-school sample were rather small, further work is needed to substantiate these results.

V.—DISCUSSION AND CONCLUSIONS.

In such investigations as that presented in this paper, it is necessary to consider the results from the inter-related standpoints of theory and practice. The experimental procedures and the theoretical results, outlined in the previous sections, must be shown to be valid in terms of both meaning and application in the setting of the proposed secondary school examinations. The experimental procedures involved in the three types of practical test were all acceptable to the teachers likely to be concerned with the examination, and it was clear that few difficulties would be met so long as due notice was given for the collection of required apparatus and other materials. Similarly, the manner of collection of the teacher's estimates presented little difficulty. The written attainment test was already an accepted part of the examination, but the scientific ability test received more sceptical consideration. Since this test was not a proposed part of the secondary school examination, its use in this investigation was accepted. Were its further use required, it would be necessary to justify its inclusion in any examination on both academic and practical grounds. Further use of school time was not available for obtaining the reliability (by retest) for any of the tests used, but in general terms, and by comparison with what is known of similar types of tests, these present forms could be assumed to be sufficiently reliable for their purpose.

The school differences shown in the preceding analysis were entirely due to the current stage of development in the proposed C.S.E. courses. Many schools had previously had little facilities for science teaching of any kind. Keeness and willingness to experiment had outstripped the availability of both apparatus and teaching staff. Progress towards a stable level of science teaching was known to differ from school to school. These large school differences could be expected to disappear within two or three years. Of greater theoretical and practical interest are the differences between the three science subjects on the three forms of practical test. While these also might be due to some extent to the newness of the situation, further work is needed to check this. It could be that the three subjects are sufficiently different in content, methods of work, and teaching presentation to disallow a common practical test. If this is so, then different forms will have to be used. There is some slight evidence of this in Table 1, where the six components of practical work are shown to differ in order of importance in each subject.

Reviewing the problem as initially posed by the teachers concerned, it will be noted that the demands of the proposed examination have been met. Three schemes of practical assessment have been devised and evaluated in appropriate conditions and with appropriate samples of pupils. Any one of these schemes was acceptable to the teachers concerned, and all have been shown to be fully operable under present school conditions. Indication has been given of the best practical test in each subject within the setting of the ordinary secondary school. At no point has the investigation proceeded without due attention to both the exigencies and the possibilities of the examination setting; all results are thus fully valid and interpretable within that situation and within the school system. The main results from this investigation may be stated as follows:

- (1) Three instruments for the assessment of practical work in science have been devised and evaluated for use at the level of the proposed C.S.E. examinations. At least one has been shown to be valid for use in each of the three science subjects proposed for inclusion in the examination.

- (2) The results of these practical tests show large differences between schools and these are shown to be the result of present background and teaching levels in the schools.
- (3) The results of these practical tests further show large differences between the subjects for any one type of test. Further work is needed to show whether the most appropriate form of practical test differs from one science subject to another.
- (4) The customary accepted theoretical relationships have been shown to hold between these practical tests, on the one hand, and teachers' estimates, scientific ability and attainment tests, on the other.

ACKNOWLEDGEMENTS.—The writers wish to express their appreciation and thanks for the co-operation given in this work by Miss D. J. Alexander and the teacher members of the L.S.S.E. Science Panel, and by a number of L.C.C. schools.

VI.—REFERENCES.

- ¹ PEEL, E. A. (1955). Psychology and the teaching of science. *Brit. J. Educ. Psychol.*, 25, 135-144.
- ² WARD, A. E. H. (1954). *A Study of Practical Tests in Physics at Sixth Form Level in a Grammar School*. M.A. Thesis, London.
- ³ LOMAS, W. H. (1956). *A Study of Laboratory Tests in Biological Subjects in Preparation for Advanced Level G.C.E.* M.A. Thesis, London.
- ⁴ KERR, J. F. (1963). *Practical Work in School Science*. Leicester: The University Press.
- ⁵ ABOUSEIF, A. A. (1963). *An Evaluation of Certain Practical Examinations in Science Subjects for the Prosposed Certificate of Secondary Education Examinations*. M.A. Thesis, London.

(Manuscript received 26th August, 1963.)

TEACHING METHOD AND RIGIDITY IN PROBLEM SOLVING

By M. L. KELLMER PRINGLE* AND I. R. MCKENZIE†

(*Department of Child Study, University of Birmingham*).

SUMMARY. Following the work of Luchins, a test suitable for 11-year-olds was designed to explore rigidity in problem solving among pupils in two junior schools. One of these is based on a child-centred, progressive regime while the other believes in a mainly adult-directed, traditional form of organisation. The total sample consisted of 211 children. Intelligence quotients and the results of two arithmetic tests were available for every child.

No over-all difference in rigidity was found between the two schools. However, within each school, children in the lowest stream and pupils of more limited intelligence showed a significantly higher degree of rigidity. Moreover, in the progressive school, the less able children showed greater flexibility in problem solving than the comparable groups in the school using traditional teaching methods. Thus, it seems, that intelligence and educational approach have a differential effect on rigidity, rather than an "all or none" influence as has been suggested previously.

I.—INTRODUCTION.

In current psychological literature, the terms 'set' or 'Einstellung' are used to denote rigidity of thinking in problem solving. Previous experience, repetition and type of training are among the functional determinants of perception, operating in problem solving and affecting the degree of rigidity. That structural determinants are likely to play an important part has been suggested by work with brain injured, feeble minded, psychotic, deaf and blind patients (Lewin, 1935; Cameron, 1939; and Luchins, 1942). It has also been demonstrated that frustration or stress may lead to an increase in rigidity (Beier, 1951; Cowen, 1952; and Gynther, 1957).

Luchins (1942) suggested that teaching method might be another variable influencing the degree of rigidity exhibited in children's thinking. His view was based on a comparative study of pupils attending, respectively, a private and public school (i.e., American State school). Neither level of intelligence nor age were found to be significant variables. Though he argued that "the method of teaching in the private schools was of a more progressive nature than that used in the public schools", his sample of private school children was too small to justify statistical comparisons. Moreover, the difference found could have been due to selective, environmental factors: children attending private schools tend to be drawn from sections of the population which are economically, socially and culturally privileged and usually of higher intelligence. Thus, Luchins's conclusions seem to be based more on inference than on statistical evaluation and to omit a consideration of socio-economic influences.

Rokeach (1950) defined rigidity as "the inability to restructure a field in which there are alternative solutions to a problem in order to solve that problem more efficiently". He found that high-prejudiced children and adults—as defined by their responses to a prejudice scale—were significantly more rigid in

* Now National Bureau for Co-operation in Child Care.

† Now South Shields Child Guidance Service.

solving various non-social problems than low prejudiced subjects. In another experiment, Rokeach (1950) tested the effect of perception time upon rigidity of thinking. His hypothesis was that "The person in a hurry will be forced to perceive a problem confronting him narrowly and, simultaneously, will be forced to resort to behavioural supports of a concrete nature in order to solve the problem. One consequence of having or not having enough time for perceiving-thinking is flexibility or rigidity of behaviour." The results confirmed his expectation that rigidity would decrease with increased perception time, except for an important proviso: there seemed to be an optimal time allowance for each task which gave maximum returns but further delay did not result in increased flexibility.

The main aim of our experiment reported here was to explore whether teaching method is an operative variable in determining the degree of rigidity in a Luchins's type problem solving situation. Werner (1946) has pointed out the ambiguity of the term rigidity; whether it should be defined functionally or structurally and whether it should be considered a uniform or a multiform trait. While recognising these problems one would expect that those general personality factors, likely to affect rigidity, would be randomly scattered in the two schools used in this study. In our experiment, rigidity is considered in terms of ability 'to break set' in problem solving activities. The hypothesis to be tested was that a progressive teaching approach would lead to a reduction in rigidity. The influence of intelligence and sex was also considered.

II.—EXPERIMENTAL DESIGN AND TESTING TECHNIQUES.

(1) *The subjects.*

This investigation forms part of a larger, longitudinal project being carried out by one of the authors (M.L.K.P.). It is concerned with various aspects of children's intellectual, emotional and educational development during the junior school years. The 211 children, who took part in this study, comprised the entire fourth year in two junior schools in the Midlands; there were 101 boys and 110 girls. Both schools are situated in areas populated mostly by skilled working class families, many of whom live in owner-occupied houses; the proportion of professional, managerial and business people is small as is the number of semi-skilled and unskilled workers; a fair proportion of families live in Council houses, some of which are on post-war estates.

The two schools which the children attend differ markedly regarding teaching approach and school organisation. One school (referred to as school T) considers attainments in the basic subjects to be of utmost importance; makes regular and frequent use of standardised tests on the basis of which the pupils are streamed; and teaching methods and time-tabling follow a traditional, adult-controlled pattern. The other school (referred to as school P) considers the all-round development of each child to be of utmost importance; the competitive element which is so much to the forefront in school T is absent and there is no streaming until the fourth year; much emphasis is placed on group work and on co-operative ventures, which may extend over weeks or months; teaching is based not on separate subjects but on projects in which art, craft, drama, music, nature study and the like all play a part; the curriculum and discipline are largely child-centred. Both schools are happy communities and are excellent of their kind, being given leadership and direction by devoted head masters and staffs in which very few changes occur, teachers leaving mainly for promotion.

(2) *The Experimental Design.*

(i) *The Pilot Experiment.* Here the subjects were forty children, with a range of ability similar to that of the main experimental sample. The purpose of the pilot experiment was threefold :

- (a) To establish a suitable level of difficulty in number work for 11-year-olds.
- (b) To establish a suitable time limit to be allowed for the completion of each problem.
- (c) To obtain insight into the usefulness, clarity and effectiveness of the proposed experimental instructions ; also to note the timing of these.

Since only four children failed, the level of difficulty seemed suitable for our age group. Each item was comfortably completed in 60 seconds and, therefore, 90 seconds were allowed for each item in the main experiment so that even the slowest child would not have to hurry. The final instructions aimed to overcome procedural weaknesses revealed in the pilot run.

(ii) *The comparability of the two groups of children comprising the main sample.*

Since the arithmetical attainment needed for the experiment was only at about the 7-year-old level, it was likely that almost all the pupils in both groups would be able to satisfy the required criteria of arithmetical attainment. Since Luchins did not find intelligence to have a significant effect on rigidity, it was not essential that the two groups should be matched for ability. However, it was felt to be advisable to have available some information on the children's arithmetical attainments. For this purpose the results of two group tests were used, namely, the N.F.E.R. Arithmetic Test 7 and the Vernon Arithmetic-Mathematics Test (Buros, 1957). The former was chosen as being the most suitable for our age group while the latter was likely to give more 'head room' for the ablest children. With regard to intelligence, Terman-Merrill I.Q.s. were available for all the subjects (Form L, 1937).

(iii) *Testing Rigidity.*

A test to assess degree of rigidity was devised by one of the authors (I.R.M.). The problems, used by Luchins, were simplified to make them more suitable for the children in our sample. Smaller numbers were introduced to make the actual arithmetical calculations simpler and the solution was changed from B-A-2C to B-A-C. Though the problems resemble the Ingenuity Problems on the Terman-Merrill Intelligence Scale, Form L (1937) at year XIV, the questions are very much easier. In addition, greater help is initially given since the necessary thought processes are demonstrated by working through three examples with the children. The test itself consists of eleven items.* Of these, the first three aim at establishing 'set', the third being solvable only by the 'set' solution ; the next two (4 and 5) can be solved by the 'set' process or by a simpler alternative ; problem 6 is a 'set' item while problems 7 and 8 are again solvable either by the 'set' procedure or by a simpler alternative ; problem 9 cannot be solved by the 'set' formula : thus, the number of failures on this problem, compared with the number of failures on the other ten problems, gives some measure as to the disrupting effect of 'set' ; lastly, problems 10 and 11 (like problems 4, 5, 7 and 8) can be solved both by the 'set' procedure and by a simpler alternative.

* Copies of the test, together with the key and instructions for administering it, can be obtained from Adam House, 1, Fitzroy Square, London, W.1.

(3) *Treatment of the Results.*

In the pilot experiment it was found that the degree of 'set' or rigidity varied from one child to another. Since results were not 'all or none', a system of scoring was devised to classify the relative strength or intensity of rigidity shown by each child. Four categories were distinguished and the criteria for each were as follows:

- (a) 'Set' of the 1st order: the subject fails item 9 and returns to the 'set' solution for both items 10 and 11.
- (b) 'Set' of the 2nd order: here there are two criteria:
 - (i) item 9 is failed and the subject then discontinues the use of the 'set' solution for the last two items, passing them by the direct method, or,
 - (ii) the subject passes item 9 by a direct method but returns to the 'set' solution for items 10 and 11.
- (c) 'Set' of the 3rd order: the subject uses the 'set' solution up to item 9 but discards it with this problem. The direct method is used for items 9, 10 and 11.
- (d) No 'set': the subject does not use the 'set' solution for item 9 and makes use of it *no more than twice* after problem 3.
- (e) Misunderstood: into this category were placed subjects who had obviously been unable to cope with the problems, be it because of limited ability, lack of arithmetical skills or a combination of both.

III.—RESULTS.

(1) *Arithmetical Attainment and Intelligence.*

The arithmetical attainment of the children in school T was found to be significantly higher than that of school P pupils on both the tests used. There was no difference, however, regarding the variance on either test (Table 1). There was also a significant difference with respect to intelligence.

TABLE 1

ARITHMETICAL ATTAINMENT AND INTELLIGENCE FOR EACH SCHOOL.

(P value of significance of differences from T test.)

	N	N.F.E.R. Arith. T.7			Vernon Arith. Math ^o T.			Terman-Merrill Int.T.		
		Mean	σ	P	Mean	σ	P	Mean	σ	P
School T	109	104.09	12.80	< .01	35.84	9.05	< .01	116	18.54	< .01
School P	102	97.72	12.59		31.75	9.70		108	17.20	

Inspection shows that the four distributions are normal but as a check the least normal looking one was treated by the Chi-Square One Sample Test. This was found not to be significant.

The experiment might have been neater, had the arithmetical attainment and intelligence levels of the two groups of children been closely similar. However, the fact that this was not the case, does not invalidate the findings; indeed, it need have little bearing upon them. What had to be established was that the great majority of children in each school had obtained a level of

arithmetical skill well beyond that necessary for the problem solving tasks of the rigidity test. The same applied to level of intelligence. Moreover, the intelligence variable could to some extent be controlled by comparing those children in each school who fell within certain ranges of ability.

2) *Degree of Rigidity found in each school.*

Just over half the subjects showed a considerable degree of 'set' or rigidity (1st order 'set') while only 12 per cent. showed 'no set.' The number of children who failed to grasp what was required was very small ($2\frac{1}{2}$ per cent.) and they were omitted from subsequent statistical evaluation; thus, the total working sample consisted of 206 children, 98 boys and 108 girls. Because the number showing 3rd order 'set' was also too small to constitute cell groups of their own, they were combined with the 'no set' category in the computation of chi-squares (Tables 2 and 3).

TABLE 2
DEGREE OF RIGIDITY IN EACH SCHOOL.

	N	1st Order	2nd Order	3rd Order	No 'set'	Misunderstood
School T ..	102	54	29	4	14	1
School P ..	109	57	34	3	11	4

TABLE 3
DEGREE OF RIGIDITY BY SCHOOL AND SEX.

	N	1st Order	2nd Order	No 'set'
School T	101	54	29	18
School P	105	57	34	14
Both	206	111	63	32
All Boys	98	51	30	17
All Girls	108	60	33	15
School T Boys	48	27	13	8
School T Girls	53	27	16	10
School P Boys	50	24	17	9
School P Girls	55	33	17	5

None of the differences between the schools was found to be significant. The two groups of children were similar with regard to rigidity (i.e., proportions showing 'complete' or 1st order set); with regard to flexibility (i.e., proportions showing 'no set'); and regarding the numbers who failed to solve problem number 9. Thus, the hypothesis, that the school using progressive methods would show less rigidity than the more formal school, must be rejected.

(3) *Sex Differences.*

Here also children who failed to understand the task were omitted and those showing 3rd order 'set' were combined with those showing 'no set'. None of the sex differences, either within each of the schools or between the sexes combined, proved to be significant (Table 3).

(4) *Intelligence.*

To explore the relationship between rigidity and intelligence, five methods were used:

(a) The results of the rigidity test were considered in relation to streaming (there were three streams in school T and four in school P). Although streaming is not an unreasonable measure of the relationship between rigidity and ability, two points have to be borne in mind: firstly, that streaming is a classification based mainly on a pupil's *educational* attainment rather than on his intelligence; and secondly, that a teacher-variable is also being introduced.

In both schools the lower streams were found to show a significantly higher degree of rigidity with one exception: this was the lowest stream (D) in School P, where this trend was reversed and the children showed, in fact, less rigidity than streams B or C.

A chi-square trend analysis (Maxwell, 1961) showed a marked departure from linearity in school P; however, when class D was excluded from the calculation, a very similar pattern of linearity was found in both schools (Tables 4 and 5).

TABLE 4
A COMPARISON OF RIGIDITY AND STREAMING: OBSERVED FREQUENCIES OF DIFFERENT LEVELS OF SET.

(Expected frequencies shown in brackets.)

Stream	School T				School P			
	N	1st Order	2nd Order	3rd Order and no set	N	1st Order	2nd Order	3rd Order and no set
A	39	15 (21)	15 (11)	9 (7)	25	5 (13.5)	16 (8)	4 (3.5)
B	38	18 (20)	11 (11)	9 (7)	26	15 (14)	8 (8.5)	3 (3.5)
C	24	21 (13)	3 (7)	0 (4)	29	26 (16)	2 (9.5)	1 (3.5)
D	—	—	—	—	25	11 (13.5)	8 (8)	6 (3.5)

TABLE 5
ANALYSIS OF THE ASSOCIATION BETWEEN STREAM AND LEVEL OF SET WITHIN SCHOOL T AND SCHOOL P*

Source of Variation	Degrees of freedom	School T χ^2 P	School P (3) χ^2 P	Degrees of freedom	School P (4) χ^2 P
Linear regression on streams	1	11.61 < 0.01	20.2 < 0.01	1	2.2 > 0.05
Departure from linearity	3	4.09 > 0.05	7.1 > 0.05	5	27.59 < 0.01
Overall Chi-square value	4	15.7 < 0.01	27.3 < 0.01	6	29.79 < 0.01

* In School P only this has been calculated for the top three and for all four streams, respectively.

(b) The I.Qs. obtained from the Terman-Merrill Scale were arranged in four groups, ranging from dull to superior ability (Tables 6 and 7). School T

TABLE 6
INTELLIGENCE AND RIGIDITY

I.Q. range	School T—Level of set				School P—Level of set				Total Sample—Level of set			
	N	1st	2nd	3rd and no set	N	1st	2nd	3rd and no set	N	1st	2nd	3rd and no set
130+	28	8	12	8	11	4	4	3	39	12	16	11
110-129	35	17	11	7	38	16	17	5	73	33	28	12
90-109	30	21	6	3	41	26	10	5	71	47	16	8
70-89	8	8	0	0	15	11	3	1	23	19	3	1

TABLE 7
ANALYSIS OF THE ASSOCIATION BETWEEN INTELLIGENCE AND LEVEL OF SET WITHIN
SCHOOL T AND SCHOOL P

Source of Variation	Degrees of freedom	School T χ^2 P	School P χ^2 P	Total Sample χ^2 P
Linear regression on streams	1	14.96<0.01	6.03<0.05	20.6<0.01
Departure from line- arity	5	3.95>0.05	5.92>0.05	2.2>0.05
Overall Chi-square Value	6	18.91<0.01	11.95>0.05	22.8<0.01

had nearly three times as many children with I.Qs. of 130 or more, but only half as many in the lowest category (I.Q. range 70 to 89), when compared with school P. While a significant relationship between intelligence and rigidity was found both for the total sample and for school T (beyond the 1 per cent level), in school P this did not quite reach the 5 per cent level. There was evidence of linear regression with respect to the total sample within school (again beyond the 1 per cent level) while in school P significance reached the 5 per cent level but not the 1 per cent level).

(c) The mean I.Q. of those passing problem 9, the most crucial item of the rigidity test, was compared with that of the group who failed it. Though the correlation was not high, the Point Bi-serial Coefficient did indicate a definite relationship (Table 8).

TABLE 8

INTELLIGENCE AND PERFORMANCE ON PROBLEM 9 OF THE RIGIDITY TEST.

	N	Mean I.Q.	σ	r	t	P	t(Mean)	P
Passing Problem 9	94	118.5	16.32	+.36	5.42	<.01	18.5	<.01
Failing Problem 9	112	106.1	16.34					

(d) The difference between the mean I.Q. of the group which successfully managed problem 9, was found to be highly significant compared with that of the group who failed it (Table 8).

(e) Since intelligence appeared to be an operative variable, it was decided to explore whether the two different teaching methods were perhaps having a differential influence on rigidity in problem solving with respect to different ability groups. Again, success or failure on problem 9 was taken as the criterion. The results suggest that progressive teaching methods may result in a lower degree of rigidity where children of low average or less ability are concerned; but for bright or very able children teaching methods do not appear to have any effect in this respect (Table 9).

TABLE 9

PERFORMANCE ON PROBLEM 9 IN RELATION TO INTELLIGENCE LEVEL IN EACH SCHOOL.

I.Q.	School	N	Pass	Fail	Significance
121 and above	T	44	29	15	$\chi^2 = 0.29$ $P > .05$
	P	24	17	7	
	Both	68	46	22	
101-120	T	37	17	20	$\chi^2 = 0.83$ $P > .05$
	P	44	16	28	
	Both	81	33	48	
100 and less	T	20	1	19	$\chi^2 = 4.95$ $P < 0.05$
	P	37	14	23	
	Both	57	15	42	

IV.—DISCUSSION AND CONCLUSIONS.

This experiment has to be regarded as a pilot study since only two schools took part; moreover, the test used to assess rigidity in problem solving was specially devised for it and no standardisation data were available. Thus, any conclusions must be regarded as being highly tentative.

To judge by the small number of children who misunderstood or failed to grasp what had to be done, it seems that this modified Luchins's rigidity test is of a suitable level of difficulty for 11 year olds. At this age only a small minority were found to show no rigidity while over half the children exhibited a considerable degree of 'set'.

With regard to the influence of educational method, no over-all difference in rigidity was found between the two schools. Nor were there any sex differences within or between the schools. To this extent, then, our findings do not bear out the hypothesis put forward by Luchins. On the other hand his view, that intelligence does not influence rigidity of thinking, seemed to derive some support from the fact that, although the level of ability was significantly higher in the traditional than in the progressive school, this did not influence the overall degree of rigidity. However, further analysis revealed that neither intelligence nor educational method exert an 'all or none' influence but that they have a more subtle, differential effect.

Firstly, when the sample was grouped according to level of intelligence, the lowest showed a significantly higher degree of rigidity. This was the case for the total sample, both on the whole test and on its most crucial problem (No. 9). It suggests that limited intellectual ability has a detrimental effect, leading to reduced flexibility in problem solving. One would perhaps have predicted such a finding in the light of general psychological theory regarding cognitive ability.

However, the pattern in school P does not entirely fit this conclusion. This leads to a reconsideration of the influence of teaching methods. Here interpretation presents considerable difficulties. On first inspection, it appears that the progressive methods practised in school P have a beneficial influence among the least able children leading to greater flexibility and reduced rigidity when compared with a more traditional educational approach; the overall relationship between intelligence and rigidity is not significant in the progressive school; the most crucial item (No. 9) is passed by a significantly greater proportion of school P than school T pupils with I.Qs. of 100 or less; moreover, the tendency for lower streams to show more rigidity (apparent in both schools) is quite reversed in its lowest stream. Of course, this last conclusion rests on the assumption that the lowest stream in each school is comparable.

There are two facts, however, which must be set against this interpretation: firstly, there were more intellectually dull and educationally backward children in the progressive school; secondly, there was no suggestion of this reversed trend in the preceding stream (i.e., stream C in school P). Thus, an alternative explanatory hypothesis must be considered. This is that the D stream in school P has no real counterpart in school T and that its pupils are so dull and backward that they have not even acquired 'set'; underlying this assumption is the view that rigidity is itself learned and that it presupposes a level of mental maturity which enables a child to generalise and then consistently to apply this generalisation.

If the former hypothesis is accepted, one would have to argue that a progressive school regime has a beneficial effect on rigidity only for the dullest children. It still leaves open to question whether and to what extent this is due to the actual teaching methods or to the influence of a child-centred approach on a pupil's intellectual and emotional development. It is worth recalling here that rigid streaming was practised throughout all four junior school years in school T but introduced in school P only in the fourth year. Even then the obvious order of merit, implied by using the letter A, B, C, etc., was avoided; each class was known by the initials of its teacher's surname (a practice which applied also throughout the preceding years). It could be that a less competitive school environment reduces feelings of frustration and stress among the least able pupils and that is this which leads to reduced rigidity in thinking. Previous findings lend some support to such a hypothesis; but

further study is clearly needed with larger numbers of more adequately matched groups to determine the influence of teaching methods on rigidity among children of limited ability.

At a later stage, it is planned to consider the findings of this investigation in relation to all the other data of the longitudinal study. It will then be possible to explore also what relationships, if any, exist between rigidity in problem solving and non-cognitive variables, such as anxiety, persistence and emotional as well as social adjustment.

Should our two main conclusions be born out by further and larger scale studies, they would have considerable theoretical interest as well as be of great practical significance. They would, after all, have some bearing on the teaching methods most suitable—because most effective in at least one respect—for more than a quarter of the pupils in secondary modern schools.

ACKNOWLEDGEMENT. Analysis of the data was facilitated by a grant from the Nuffield Foundation.

V.—REFERENCES.

- BEIER, E. G. (1951). The effect of induced anxiety on flexibility of intellectual functioning. *Psychol. Mon.*, 65 (365), 1-20.
- BUROS, O. K. (1959). Graded arithmetic mathematics test No. 476, 599. *The Fifth Mental Measurements Year Book*. The Gryphon Press.
- CAMERON, N. (1939). Deterioration and regression in schizophrenic thinking. *J. Abn. and Social Psychol.*, 34, 265-270.
- COWAN, E. L. (1952). The influence of varying degrees of psychological stress on problem solving rigidity. *J. Abn. and Social Psychol.*, 47, 512-519.
- GYNTHER, R. A. (1957). The effects of anxiety and situational stress on communicative efficiency. *J. Abn. and Social Psychol.*, 54, 274-276.
- GUILFORD, J. P. (1950). *Fundamental Statistics in Psychology and Education*. McGraw Hill.
- LEWIN, K. (1935). *A Dynamic Theory of Personality*. McGraw Hill.
- LUCHINS, A. S. (1942). Mechanisation in problem solving—the effect of Einstellung. *Psychol. Mon.* 54, No. 6.
- MAXWELL, A. E. (1961). *Analysing Qualitative Data*. Methuen.
- ROKEACH, M. (1950). Rigidity and concreteness in thinking. *J. Exp. Psychol.* 40, No. 2, April, 1950, 206.
- WERNER, E. (1946). The concept of rigidity. A critical evaluation. *Psychol. Rev.*, 1946, 53, 43-52.

(Manuscript received 11th May, 1964.)

LEVELS OF PERCEPTUAL ORGANIZATION AND OF PERFORMANCE AFTER TIME FOR REFLECTION

By JOSEPHINE KLEIN
(Nuffield College, Oxford)

SUMMARY. Children between the ages of 11 and 15 marked Raven's *Progressive Matrices* twice: first at speed, so that the test resembled one in perceptual organisation; later at leisure, to allow time for thought. It was predicted that more younger children and more educationally backward children would, on the second occasion, change right answers to wrong ones. The suggestion is that, at the lower levels, language-handicaps affect reasoning-ability adversely.

I.—BACKGROUND.

THERE is some evidence now which suggests that in the lower social strata there is a greater discrepancy between performance on verbal intelligence tests and non-verbal ones, and that this discrepancy is greater for children with a higher total intelligence score than for children with a lower total score (Bernstein, 1961; Daniels, 1964). It is possible that some cleverer but socially less favoured children have a language handicap.

It is established that linguistic performance improves with age (e.g., Lawton, 1964). The capacity to 'reason' also improves with age. Within limits, the child will be able to solve logical problems at a later age which were baffling him at an earlier stage.

There is as yet no commonly agreed conceptual framework in psychology which would give an acceptable theoretical formulation of the connection between these findings. Rather than make explicit a theory into which the experiment described in this paper might fit, the present investigator's procedure and findings are set out with the minimum of conceptualisation. The prediction was made from a hunch, which has been gratifyingly strengthened by the findings.

II.—EXPERIMENT.

The experiment was conducted in a secondary modern school in a small town in Oxfordshire.* Children in five classes were tested, two classes being for backward children. Not all children were tested, because of absence at one or other session. Copies of Raven's *Standard Progressive Matrices* were handed out, and the children instructed in their use in the normal way.† The teacher then said:

"Draw a circle round the right one. See how many you can get right. Each set begins with some easy ones. So try to put down an answer for each. They get harder, but try to put down an answer anyway. If you can't see the answer straight away, draw a circle round the one that seems most likely and don't worry. Try to get as near to the end of the book as you can. Tomorrow (or some other day) we will look at them again and you can make changes if you want to. Today I want you to work as fast as you can. There is a test on each side of the page. Remember to do both sides."

*The experiment was conducted by Mrs. Coral Wabe and some colleagues, whose help I gratefully acknowledge.

†This expensive procedure was made possible by a grant from Nuffield College.

On the next occasion, the teacher said :

" Today you will have a little more time to look at what you did before. You may have changed your mind on some, since then. If you have, you can show it by putting a circle round the answer you now think right. Use your coloured pencil.

" If you *haven't* changed your mind, put a tick on the page to show that your last answer was the right one. Don't forget to tick the right one. So every page you have written on must have either a tick or a new circle. Are there any questions? (Pause.) Today there is no hurry. Take as much time as you like."

The children worked in an orderly fashion ; there were only one or two cases in which a page was omitted by accident. Most children got to the end of the series, though, of course, not with answers all correct.

III.—PREDICTION AND FINDINGS.

It was expected that children would make changes on the second occasion. They could change from wrong to right, from one wrong answer to another, and from a right to a wrong one.

It was predicted that there would be a greater number of children making changes from right to wrong, at the lower ages, and at the lower levels of educational attainment. With one exception, this prediction was confirmed. The probability that five classes would arrange themselves by chance in perfect order or at most with one transposition is $1/4 := 1 : 24$, significant at a 0.05 level.

TABLE 1
CORRECTIONS FROM RIGHT TO WRONG IN THE FIVE SCHOOL CLASSES.

	Total number of children	Ages	Number correcting from right to wrong	[Proportion of children in class correcting from right to wrong]
3Y ('normal') ..	24	13—14	5	.21
2Y ('normal') ..	24	12—13	6	.25
1Y ('normal') ..	26	11—13	8	.31
3M ('backward') ..	10	13—15	4	.40
1M ('backward')	11	11—13	4	.36

In commonsense language, it may be that the younger and less clever children are more prone to confuse themselves when they have time to think—the investigator's hunch is that they have not the linguistic apparatus which would enable them to think out *why* an answer is right or wrong, and that this accounts, at least in part, for the occasional change from right to wrong. It is possible that, given help in accurate speech, this handicap could be overcome and that then their performance on tests like Raven's *Progressive Matrices*, administered in the normal fashion, would improve.

It is necessary thus tentatively to outline what lay behind the hunch, because an objection could be raised, at this level of discourse, which it is possible to refute from the present data. Might it not be that the younger and less clever children get more worried when faced by what they have done, having in the past had plenty of experience of being wrong and needing correction? If this were so, one should expect more corrections of all kinds as one goes down the scale. But this is not so.

TABLE 2

CORRECTIONS AND RIGHT ANSWERS; AVERAGES OF THE TOTAL CLASS SCORES.

	3Y (24)	2Y (24)	1Y (26)	3M (10)	1M (11)
Wrong to wrong	1.6	2.1*	1.6	2.7	3.8
Wrong to right	3.2	4.7	3.4	3.2	2.0
Right	44.7	36.5	38.4	40.3	27.7

TABLE 3

CORRECTIONS AND RIGHT ANSWERS; AVERAGES OF CHILDREN WHO MADE RIGHT-TO-WRONG CORRECTIONS.

	3Y (5)	2Y (6)	1Y (8)	3M (4)	1M (4)
Wrong to wrong	3.0	3.0*	1.5	3.0	5.5
Wrong to right	4.8	7.3	2.6	3.2	1.7
Right	44.6	33.3	38.7	32.5	24.0

* A girl who made 22 wrong-to-wrong corrections has been excluded from the average.

In summary, the only at all convincingly regular pattern which emerges is the one based on a hunch that time for thought leads younger and less clever children falsely to modify their correct answers, which they had arrived at by means of a procedure which strongly encouraged reliance on perceptual organisation. That an important variable may be at stake here is suggested by the fact that it was brought to light even by so crude an experiment as the one here described. The evidence is not strong, statistically, but is enough to make replication on a larger scale desirable.

IV.—REFERENCES.

- BERNSTEIN, B. (1961). Social class and linguistic development. *Education, Economy and Society*, edited by Halsey, A. H., Floud, J., and Anderson, C. A. Free Press, Glencoe, Ill.
- DANIELS, J. (1964). Social class and verbal intelligence. *New Society*. 21st May, 1964.
- LAWTON, D. (1963). Social class differences in language development. *Language and Speech*, VI.

(Manuscript received 28th May, 1964.)

UNSETTLEDNESS, MALADJUSTMENT AND READING FAILURE: A VILLAGE STUDY

BY ROBIN E. GREGORY

(Senior Lecturer in Education, Matlock Training College)

SUMMARY. Use of the Bristol Social Adjustment Guides in the only primary school of a West Berkshire village revealed a significant connection between reading failure and Restlessness throughout the school, and between reading failure and Anxiety for the Approval of Other Children among the 9, 10 and 11-year-olds. Since there was apparently no tendency for Restlessness to increase with age, it seemed possible that such a form of unsettledness or maladjustment (which Stott defines as "an inability to persevere, concentrate or reflect, and a liking for easy moment-to-moment satisfaction") had been a contributory cause of reading failure. It appeared that Anxiety for the Approval of Other Children (sometimes, as Stott points out, to the extent of being led into mischief) might possibly have been caused, at least in part, by reading failure. Although the numbers of children tested were such that there remained some doubt concerning the significance of the conclusions, the results suggested that the incidence of a "Restlessness: Reading failure: Anxiety" sequence in children would justify further investigation.

I.—INTRODUCTION.

A discussion of previous research.

WITH the exception of Bennett (1938) most research into the relationship between emotional difficulties and personality maladjustment on the one hand, and reading disability on the other, agrees that there is a significant correlation. Writers have not agreed, however, either in defining the precise relationship between specific forms of maladjustment and reading or in establishing whether maladjustment is a cause, an effect or merely a concomitant factor of reading disability. At one extreme is Gann (1945), who stated that every personality tension unfavourable to reading had arisen before a child entered school. At the other are Fernald (1943) (who considered that seventy four out of seventy eight cases of emotional maladjustment were caused by reading failure), and Wilking (1941) (who similarly considered twenty nine cases out of thirty). That personality maladjustment and emotional difficulties can cause reading failure has been asserted by Blanchard (1928), Challman (1939), Gates (1941), Mary Vera (1942), Sylvester and Kunst (1943), Gann (1945), Robinson (1946), Stewart (1947), Malmquist (1958), and Pringle and Sutcliffe (1960). That maladjustment can be caused by reading failure has been asserted by Blanchard (1928), Orton (1937), Challman (1939), Preston (1940), Monroe (1946), Schonell (1948), Solomon (see Robinson, 1953), and Malmquist (1958).

'Result or cause?' arguments may lead to insoluble 'chicken egg' problems: sometimes such a pitfall has been avoided by statements that reading failure is 'related to' or 'correlated with' emotional maladjustment, statements which conveniently ignore questions of cause and effect. It is pertinent to ask why questions of priority and causality can produce apparent disagreement. Vernon (1957) gives it as her opinion that a great many of the studies of the relationship between emotional maladjustment and reading difficulty have failed to differentiate mild cases of retardation (which often show little evidence of

maladjustment) from severe cases of real failure to learn to read. In addition, many workers have been hampered by the absence of any reliable, objective measures of maladjustment, and have been forced to have recourse to teachers' opinions, clinical reports and the like.

For practical purposes, it seems useful to define emotional maladjustment as "abnormal or defective behaviour which may reasonably be assumed to have an emotional basis." When *all* recognisable forms of maladjustment are considered together, an overall negative correlation with reading ability is usually noted. When, however, a number of specific forms of maladjustment only are considered, a zero or even positive correlation may by chance result, since the relationship between some of these manifestations of emotional disturbance and reading ability may be positive. One could well understand a withdrawn child, for instance, reading well because of a desire to escape from the world into books. Indeed, Haggard (1957) demonstrated that in certain circumstances a maladjusted child can have a particularly high reading achievement.

The present study.

Bristol Social Adjustment Guides and Adjustment Pointers were completed by the respective class teachers for fifty three children in the only school in a small West Berkshire village. The children were in two distinct groups:

- (i) every child in the school who was aged 8-10 or over ($n=35$)
 - (ii) every child in the school who was aged between 6-0 and 7-10 ($n=18$).
- It was hoped that by having a 'younger' as well as an 'older' group, some useful comparative data would be obtained.

The 'older' group was tested for mechanical reading ability with the Holborn Reading Scale, and for intelligence with either the Sleight or the Pidgeon Non Verbal test, or both. Using these R.As. and I.Qs., the thirty five children were arranged into three categories: non readers, severely retarded readers, and others. A severely-retarded reader was defined as one whose R.A. was at least two years below his chronological age, for which discrepancy intelligence could not be regarded as a cause. In practice it was not found necessary to select a particular I.Q. which disqualified a child from the severely-retarded category since examination of the test results showed at once which children were severely below normal intelligence.

Children in the 'younger' group (ii above) were placed in two categories only: poor readers, and normal or above average readers. Assignment to either of these categories was determined by performance on the Holborn Reading Scale reinforced by the teacher's opinion. Inevitably, the results for these younger children were less precise than those for the older, but it was hoped that they might provide useful corroborative or comparative data.

II.—RESULTS.

Adjustment Pointers.

Completion of the Adjustment Pointers gave the results shown in Table 1, where the figures are compared with those for two types of school described by Stott: a 'normal' junior school, and "a backward or other class having a high proportion of children whose unsettledness and maladjustment arise mainly from conditions outside the school" (Stott's percentage figures are derived from column ii, page 9, of the *Manual to the Guides*).

TABLE 1

ADVERSE ANSWERS TO POINTERS: VILLAGE CHILDREN COMPARED.

	Adverse answers to Pointers						
	0	1	2	3	4	5	6
No. of village children	32	10	6	3	2	0	0
Village Percentage	60.4	18.8	11.3	5.7	3.8	0	0
Stott's 'normal' Percentage ..	71.2	10.5	6.2	5.5	4.5	1.3	0.6
Stott's 'maladjustment outside the school' Percentage	41	16	11	15	14	2	1

Compared with Stott's 'normal' percentages, it appeared that rather fewer of the fifty-three village children showed no adverse answers to Pointers and rather more showed one or two adverse answers. The village school seemed to fall midway between the two types of school described by Stott. He states, however, that as far as the environment provided by the school is concerned, the tell-tale indication is not the proportion of unsettled children, but the distribution of the severity of unsettledness among those who have any tendency to be so: "In a good school situation there will be, among the unsettled or unresponsive, more who are slightly so—that is, recording one adverse pointer only—and a steadily decreasing proportion with higher adverse marks." It may thus fairly be concluded that the extent of maladjustment in the village school was not seriously greater than that obtaining elsewhere. Rather more children showed unsettledness or maladjustment than normal, but they showed it to a limited extent, and insofar as they showed it at all it was probably a result of home rather than school conditions.

Social Adjustment Guides.

Since there appeared to be no reason for supposing the village school situation to be itself productive of maladjustment, the Guides were used to discover whether individual cases of reading retardation were related to unsettledness or maladjustment. Table 2 gives total and mean entries on Diagnostic Forms under each heading, distinguished by Stott, for non-readers and severely-retarded readers, aged 8-10 or over. Mean entries under the respective headings are given for the twenty-one other children aged 8-10 or over for purposes of comparison.

Examination of Table 2 and of individual cases leads to the following observations:

(i) Withdrawal, Anxiety concerning Adults, Anxiety concerning Children and Restlessness were more prevalent among the fourteen severely-retarded readers and non-readers than among the other twenty-one children. Depression was less prevalent. Withdrawal, however, was confined to two children, so no useful conclusions could be drawn. Anxiety concerning Adults was marked in two cases (four entries and three entries), and may well have contributed to reading failure in those instances. The difference in incidence of XA as between the fourteen severely-retarded and the others was not, however, generally significant.

TABLE 2

NON-READERS AND TWO-YEAR RETARDED READERS AGED 8-10 OR OVER : ENTRIES IN BRISTOL SOCIAL ADJUSTMENT GUIDES.

	Total entries under								
	U	W	D	XA	HA	K	XC	HC	R
Non-readers (n=6) Mean I.Q.=77.5	8	4	0	7	1	0	4	0	10
Severely-retarded readers (n=8) Mean I.Q.=105.5	2	2	1	6	0	1	6	1	11
Mean entries for the 14 children above	0.7	0.4	0.1	0.9	0.1	0.1	0.7	0.1	1.5
Mean entries for the 21 other children aged 8-10 or over	0.5	0.0	0.4	0.4	0.1	0.2	0.0	0.1	0.1

Key :

U—Unforthcomingness.

W—Withdrawal.

D—Depression.

XA—Anxiety about adult interest and affection.

HA—Hostility to adults.

K—Unconcern for adult approval.

XC—Anxiety for approval of and acceptance by other children.

HC—Hostility to other children.

R—Restlessness.

(ii) Anxiety concerning Children (XC) was found in seven of the fourteen severely-retarded children but not once among the others. Restlessness (R) was found in ten of the fourteen cases, but in only three of the other twenty-one. Whilst it may be dubious practice to speak of Anxiety or Restlessness being present or not present in a child, it would seem legitimate, for purposes of statistical calculation, to speak of XC or R appearing or not appearing on a Diagnostic Form. Thus, Tables 3 and 4 may be drawn up.

TABLE 3

CHILDREN AGED 8-10 OR OVER ON WHOSE DIAGNOSTIC FORMS XC APPEARED.

	XC appears in	XC does not appear in
Non-readers or severely retarded	7 cases	7 cases
Other children	0 cases	21 cases

In Table 3, after applying Yates' Correction, Chi-squared = 10.17, and in Table 4, 9.43. In both tables, therefore, Chi-squared is significant at the 0.05 level. In the village school there thus seemed to be a significant connection between reading retardation on the one hand and Anxiety for the Approval of Other Children and Restlessness on the other.

TABLE 4

CHILDREN AGED 8-10 OR OVER ON WHOSE DIAGNOSTIC FORMS R APPEARED.

	R appears in	R does not appear in
Non-readers or severely-retarded	10 cases	4 cases
Other children	3 cases	18 cases

(iii) In one case reading failure was primarily due to low intelligence (Sleight I.Q. 57), but in the cases of the other severely-retarded or non-readers who exhibited symptoms of Restlessness or Anxiety for the Approval of Other Children this was not so, for their I.Qs. ranged from 80 to 111. There was, at this stage, no way of knowing whether R or XC had caused reading failure, whether reading failure had caused R or XC, or whether there was some other relationship.

Results for the eighteen younger children showed appreciable differences between poor readers on the one hand and normal or above-average readers on the other with regard to Restlessness, but not with regard to Anxiety for the Approval of Other Children. Table 5 shows R and XC entries for these children.

TABLE 5

YOUNGER CHILDREN: R AND XC ENTRIES.

	R entries		XC entries	
	Total	Mean	Total	Mean
6 and 7 year-old poor readers (n=9)	24 (9 children)	2.7	4 (3 children)	0.4
6 and 7 year-old normal and above average readers (n=9)	5 (4 children)	0.6	3 (one child)	0.3

III.—CONCLUSIONS AND DISCUSSION.

Use of the Bristol Social Adjustment Guides in the only school in this West Berkshire village suggested a significant connection between reading failure and Restlessness throughout the school, and (among the older children) between reading failure and Anxiety for the Approval of Other Children.

Since the incidence of Restlessness among the older children was no greater than among the 6 and 7-year-olds, it seemed unlikely that reading failure had produced restless symptoms, for if it had we should expect continued failure to have aggravated the symptoms. Indeed, it may well be that this particular form of unsettledness or maladjustment (which Stott defines as "an inability to persevere, concentrate or reflect, and a liking for easy moment-to-moment satisfaction") had been at least a contributory cause of the reading failure.

The relatively heavy incidence of Anxiety for the Approval of Other Children among the older severely retarded readers (seven out of fourteen exhibited symptoms, though admittedly not to any very marked extent), when contrasted with (a) the non-existence of such symptoms among children of a similar age without severe reading difficulty and (b) the somewhat lesser incidence among younger poor readers, suggests the possibility that this form of anxiety had been a product of the reading failure. A child, having demonstrated his inability to master so essential a skill as reading, would naturally be concerned lest he forfeit the approval of other children, possibly assuming himself already to have lost the approval of adults through his reading failure.

It can be objected that the technique whereby conclusions are drawn about children's development from a study of two discrete age groups, is of doubtful validity. The force of this objection must be conceded; what is needed is a longitudinal study of a single group of children, some of whom fail in reading. It is hoped that, despite this methodological objection, sufficient evidence has been presented to suggest a profitable field of enquiry.

ACKNOWLEDGEMENTS.—The interest shown and advice given by Professor M. D. Vernon of the University of Reading have proved invaluable. Mr. W. H. Clarke, the Headmaster of the school, gave freely of his time and facilities.

IV.—REFERENCES.

- BENNETT, C. C. (1938). An enquiry into the genesis of poor reading. *Teachers' Coll Contribs. to Educ.*, 755.
- BLANCHARD, P. (1928). Reading disabilities in relation to maladjustment. *Mental Hygiene*, 12, 772-88.
- CHALLMAN, R. (1939). Personality adjustment and remedial reading. *J. Except. Chil.*, 6, 7-12.
- DONALDSON, M. (1959). *Brit. J. Psychol.*, 50, 379-80.
- FERNALD, G. M. (1943). *Remedial Techniques in Basic School Subjects*. New York, McGraw-Hill.
- GANN, E. (1945). *Reading Difficulty and Personality Organisation*. New York: King's Crown Press.
- GATES, A. I. (1941). The role of personality maladjustment in reading disability. *J. Genet. Psychol.*, 59, 77-83.
- HAGGARD, E. A. (1957). Socialization, personality and achievement in gifted children. *Sch. Rev.*, 318-414.
- MALMQUIST, E. (1958). *Factors Related to Reading Disabilities in the First Grade of the Elementary School*. Stockholm: Almqvist and Wiksell.
- MARY VERA, SISTER (1942). A critical study of certain personality factors as determining elements in a remedial reading program. *Catholic Educ. Rev.*, 40, 145-161.
- MONROE, M. (1946). *Children Who Cannot Read*. University of Chicago Press.
- ORTON, S. (1937). *Reading, Writing and Speech Problems in Children*. Chapman and Hall.
- PRESTON, M. (1940). Reading failure and the child's security. *Amer. J. Orthopsychiat.*, 10, 239-252.
- PRINGLE, M. L. K., and SUTCLIFFE, B. (1960). *Remedial Education—An Experiment*. University of Birmingham.
- ROBINSON, H. M. (1946). *Why Pupils Fail in Reading*. University of Chicago Press.
- ROBINSON, H. M. (1953). Diagnosis and treatment of poor readers with vision problems. Clinical studies in reading. *Suppl. Educ. Monogr.* No. 77.
- SCHONELL, F. J. (1948). *Backwardness in the Basic Subjects*. Oliver and Boyd.
- STEWART, R. (1947). *Personality Maladjustment and Reading Achievement*. Unpub. Diss., California.
- SYLVESTER, E., and KUNST, M. S. (1943). Psychodynamic aspects of the reading problem. *Amer. J. Orthopsychiat.*, 13, 69-76.
- VERNON, M. D. (1957). *Backwardness in Reading*. Cambridge University Press.
- WILKING, S. V. (1941). Personality maladjustment as a causative factor in reading disability. *Elemen. School J.*, 42, 268-79.

(Manuscript received March, 1964.)

DEVELOPMENTAL FEATURES OF BEHAVIOUR AND PERCEPTION

I.—VISUAL AND TACTILE-KINAESTHETIC SHAPE PERCEPTION*

By GERALD H. FISHER

(Head of Department of Psychology, University of Newcastle-upon-Tyne)

SUMMARY. Two experiments are reported. They are concerned with what have been described as:

(i) The '*non-manipulation paradox*,' this being related to a stage in the development of haptic perception in which children of the age of about 2½ years, when presented with objects or shapes for tactile examination, make little or no attempt to maximise the available information relating to their spatial configurations. Paradoxically, they are, nevertheless, able to identify by touch cues alone, many objects with extremely complex spatial configurations. In this experiment the textural characteristics of a group of objects was controlled, by carving models of them in wood. The results suggest that *important cues for recognition of objects at this age are textural rather than spatial.*

(ii) The '*topological-primacy hypothesis*.' Piaget and Inhelder (1956) argue forcibly that the achievement of a spatial co-ordinate system which is essentially linear is not complete until the age of about 8 or 9 years. They interpret the findings of their haptic shape perception experiments as indicating a gradual development from using 'primitive' *topological* relationships to using *linear* or Euclidean relationships based upon a vertical—horizontal system of co-ordinates. This hypothesis was considered in a situation similar to that used by Piaget and Inhelder, but within the framework of an entirely different experimental design. The results of this investigation appear to indicate that, when a response is available equally to each of the two shape categories under consideration, the '*topological-primacy hypothesis*' requires to be replaced by one of '*linear-primacy*.'

I.—INTRODUCTION.

JEAN PIAGET and his colleagues, in Paris and Geneva, have conducted a long series of ingenious experiments with young children. Resulting from these investigations, a number of important suggestions have been made about the possible nature of both cognitive and non-cognitive processes, as they are first evidenced in infants, and as they develop subsequently throughout childhood and adolescence. Some of these suggestions have been instrumental in the development of methods of teaching and training. In collaboration with Inhelder, Piaget (1956) has paid particular attention to the developmental nature of shape perception. From their investigations into haptic perception of meaningful and abstract, two and three-dimensional shapes, i.e., the recognition of shapes by touch cues alone, they have concluded that young children start by building-up very primitive spatial relationships which are described as 'proximity,' 'separation,' 'order,' and 'enclosure.' These relationships they observe to be the 'elementary' spatial relationships of topological geometry which appear to be related to the field forces postulated by Gestalt psychologists in order to explain a variety of features of shape perception.

* This paper is the first in the series "Developmental Features of Behaviour and Perception," reporting research conducted in the Department of Psychology, University of Newcastle-upon-Tyne. This research was supported by a grant from the Department of Scientific and Industrial Research No. R.6—149(8).

Piaget and Inhelder suggest that adults assume incorrectly that spatial operations involve reference to a vertical-horizontal system of linear co-ordinates in early childhood. From the results of their haptic perception experiments, they draw the conclusion that there is rather a gradual development from using primitive 'topological' relationships to using Euclidean relationships based upon a linear co-ordinate system. The course of this development is understood as being closely parallel with that of the operational stages of cognitive development and is not complete until the age of 8 or 9 years. A number of sub-stages into which spatial development may be divided are identified. Perhaps the most important features of these stages may be described as follows:

- (i) When making the first tactile contact with shapes, the very young child, up to the age of about 2.5 years, depends for recognition upon such chance discoveries, or cues, as poking his finger through a hole or accidentally feeling an edge, indentation or point. Children of this age do not handle shapes or objects in the sense of exploring them with the finger tips and moving them around within the grasp, but grip them tightly making little or no attempt to conduct a detailed tactile examination of objects.
- ii) The abstract two-dimensional shapes recognised most easily are 'topological' rather than 'linear' in outline. The description of shapes as 'topological' is not strictly correct but, for the purposes of identification, it will be used throughout this study. Topological shapes are those with irregular and asymmetrical contours, sometimes having holes in their surfaces, such as open and intertwined rings. Linear shapes are those with rectilinear or curvilinear outlines, such as squares, triangles and circles (see Figure 2)

Each of these statements may be considered as being a hypothesis relating to stages of the development of visual and tactile-kinaesthetic shape perception in children. Accordingly, they may be described as:

- (a) The 'non-manipulation paradox.'
- (b) The 'topological-primacy hypothesis.'

II.—THE NON-MANIPULATION PARADOX.

Not only Piaget and Inhelder, but Révész (1950) also has observed that what might be described as '*passive touch*' appears to be a primary stage in the development of tactile-kinaesthetic shape perception. Révész describes this as the '*static tactile process*' in which no attempt is made to maximise the spatial information relating to the configurations of objects held in the hands. Piaget and Inhelder observed that common objects such as a pencil, key, comb and a spoon were recognised fairly easily by the youngest children in their experimental group. It seems difficult to understand, however, how these very young children were able to recognise, or identify, the complex configurations of which the shapes of even apparently simple, common objects are composed, with such a high frequency of accuracy, if cues to their profiles are afforded only by stimuli of the apparently chance nature suggested. Accordingly, consideration was given to the possibility of the mechanism underlying recognition of objects in situations of this kind being actuated by some other aspect of them rather than by features of their shapes as such.

Since it was not possible to obtain a complete list of the objects used in the studies conducted by Piaget and Inhelder, a sample of thirty common household and play objects was selected, each of which was probably to be found, seen

and perhaps handled by children, in the majority of English homes. The objects were :

- | | | |
|-----------------------|----------------------|----------------------------|
| (1) a cardboard box. | (11) a fountain pen. | (21) a child's shoe. |
| (2) a book. | (12) a teaspoon. | (22) a woolly toy. |
| (3) a pencil. | (13) a milk bottle. | (23) a tea-cup. |
| (4) a penny. | (14) a comb. | (24) a pair of scissors. |
| (5) a table-knife. | (15) a banana. | (25) an egg-cup. |
| (6) a doll's table. | (16) a fork. | (26) a tooth-brush. |
| (7) a plastic beaker. | (17) a saucer. | (27) a pipe. |
| (8) a metal thimble. | (18) a top. | (28) a hammer. |
| (9) a rubber ball. | (19) a saw. | (29) a whistle. |
| (10) a glass marble. | (20) a torch. | (30) a pair of spectacles. |

In a preliminary experiment, in which seventeen children between the ages of 2 and 6 years were allowed to handle these objects whilst being unable to see them, it was found that the majority of them could be identified by children above the age of about 3.5 years. The objects recognised most easily were the banana and the penny, most difficult being the thimble and torch.

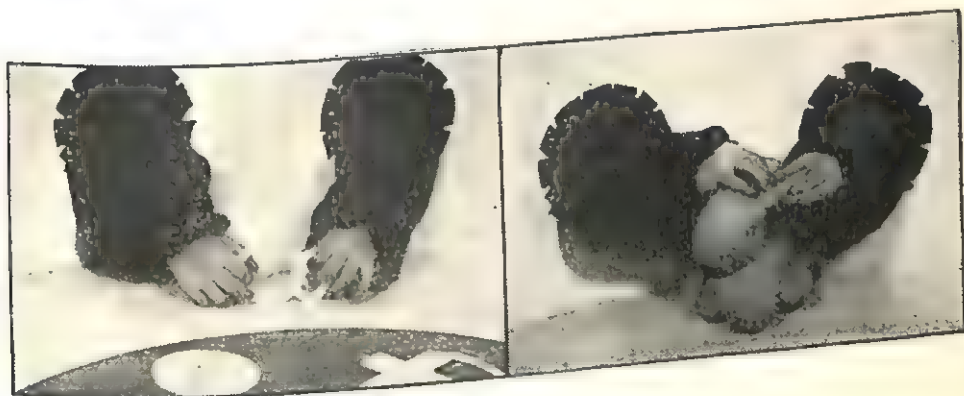


FIGURE 1

THE HANDS OF CHILDREN HANDLING THE STIMULUS MATERIALS.

In the experiments conducted with the younger Ss in this pilot study, the paradoxical nature of the stage of non-manipulation was seen most clearly. Figure 1 represents an attempt to illustrate this. The left-hand photograph shows the hands of a child of 2½ years holding a shape without exploring its configuration, while the right hand photograph shows the hands of a child of 3½ years making a detailed tactile examination of an object.

Being satisfied that the thirty objects selected were able to be identified by children using touch cues alone, the objects were used as models for carving out thirty wooden replicas of, as far as possible, exactly the same size, shape and weight as the originals. Each of the objects was photographed and reproduced as a full-size black and white matt print. These photographs were mounted in a wooden frame in order that they could be seen easily by a S sitting at a low desk with his hands and fore-arms placed through two holes in a semi-circular screen in such a way that, while he could feel objects placed into his hands readily, he was unable to see them.

TABLE 1
FREQUENCIES OF RECOGNITION OF OBJECTS AND MODELS.

S	Age		Frequency of recognition		Difference*
	yrs.	mths.	objects	models	
1	1	7	21	9	+12
2	1	8	32	17	+15
3	1	10	29	12	+17
4	1	10	50	13	+37
5	1	11	28	12	+16
6	1	11	30	10	+20
7	2	0	26	12	+14
8	2	2	24	10	+14
9	2	3	51	41	+10
10	2	3	74	24	+50
11	2	5	47	26	+21
12	2	5	78	49	+29
13	2	8	86	49	+37
14	2	8	81	47	+34
15	2	11	110	64	+46
16	2	11	63	32	+31
17	3	0	75	42	+33
18	3	0	83	49	+34
19	3	3	118	72	+46
20	3	3	76	43	+33
21	3	5	120	106	+14
22	3	7	120	110	+10
23	3	8	93	62	+31
24	3	8	118	101	+17
25	3	9	111	71	+40
26	3	9	107	98	+9
27	4	1	118	113	+5
28	4	2	115	110	+5
29	4	2	119	103	+16
30	4	2	91	86	+5
31	4	3	102	76	+26
32	4	3	91	96	-5
33	4	4	120	116	+4
34	4	5	110	112	-2
35	4	5	119	119	=
36	4	5	120	120	=
37	4	5	116	118	-2
38	4	8	120	120	=
39	4	9	120	120	=
40	4	9	120	120	=
41	4	9	107	102	+5
42	5	1	118	116	+2

$\bar{D}=17.35$
 $SE\bar{D}=2.38$

$t'=7.29$
 $p'<0.001$
 (One tailed test $V=41$)

* +ve objects>models.
 -ve models>objects.

Forty-two children were used in the experimental group. They fell more or less randomly in the age range 1.7—5.1 years. Each S was not instructed in a formal way since, particularly with the younger members, more importance was placed upon establishment of good rapport gained through informality. Care was, however, taken to ensure that each child was *shown* all the objects, models and photographs but he was not allowed to handle them before the experiment. Usually, the informal instructions included the following form of words:

"Let's play a game like blind-man's buff—Now put your hands through these holes and I will give you a toy (frequently one of the child's own toys was used for a few trials). Can you tell me what it is? Now let's try another . . ."

The objects and models were now placed into each S's hands in random order. Each model and object was presented on two occasions. There were thus 120 trials for each S. All Ss were allowed to use the photographs in order to identify the objects or models. The younger Ss particularly, insisted upon pointing to the pictures even if they knew the names of the objects or models. E frequently asked the older Ss to indicate the photograph of the object if he was in any way uncertain that it had been identified correctly by name.

In order that the overall pattern of recognition may be seen most clearly, the data have been entered in Table 1 for individual Ss. From this there appears to be a developmental trend in ability to identify objects, the younger S being able to recognise a few only, while the older Ss are able to recognise all, or almost all, of them. There is a developmental trend in ability to identify models also which follows a similar course. While the two trends are similar there is, however, a difference in recognition ability for thirty-four of the forty-two Ss, the frequency of recognition of objects being generally greater than that of models. The nature of this outcome is so clear that sophisticated statistical treatments of the data are almost gratuitous. The technique which may suggest itself in this situation is chi-square, in fact, the chi-square value relevant to these differences in frequencies of recognition is 311.20 ($p < 0.001$; $V=41$). Unfortunately, however, the data do not refer to independent samples since, as will readily be seen, the ability to identify objects and models is *correlated*. Hence, the assumptions underlying chi-square are not met by these data. We may, however, consider the data using a statistical procedure applicable to consideration of differences between correlated means. This consists essentially of testing the hypothesis that the differences in performance in the two situations have arisen by chance. The mean difference (\bar{D}) reaches a value of 17.35, the standard error ($SE\bar{D}$) associated with which is 2.38, this yields a value of 't' of 7.29, the probability of occurrence of which by chance is less than 0.001. It thus seems reasonable to conclude that these data support the view that the frequency of recognition of objects exceeds that of recognition of models.

The *shape* of both objects and models handled by Ss were virtually identical. What was achieved by carving the wooden replicas of the original objects was control over their *texture*. Hence, it seems clear that important cues for recognition of objects by children, particularly the younger ones, are those which afford information about the texture of objects rather than their shapes.

III.—THE TOPOLOGICAL-PRIMACY HYPOTHESIS.

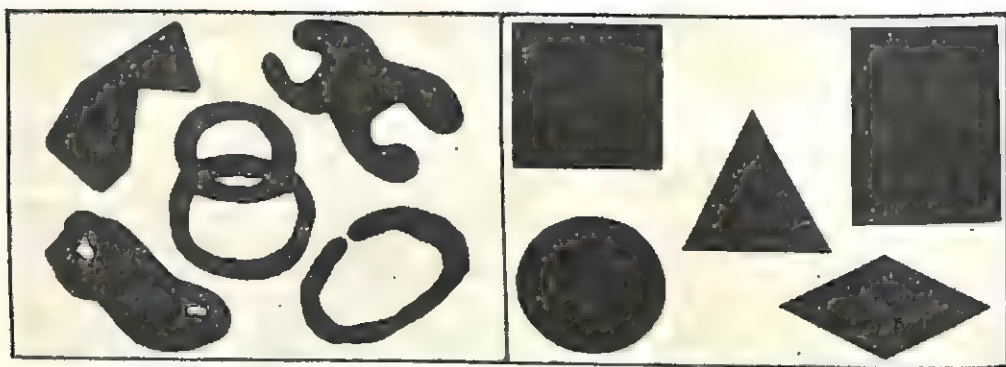


FIGURE 2

EXAMPLES OF TOPOLOGICAL AND LINEAR SHAPES.

From the results of haptic shape perception experiments, conducted in situations similar to those described above, Piaget and Inhelder argue forcibly that, when abstract two-dimensional shapes are used as stimuli, those recognised most easily have 'topological' rather than 'linear' outlines. Figure 2 shows typical examples of 'topological' and 'linear' figures. This conclusion was supported in its essentials by Page (1957 and 1959) and has recently been reaffirmed by Piaget (1962). According to Piaget and Inhelder "... *The problems set before the children face them with two distinct tasks. First, the translation of tactile-kinaesthetic perceptions into visual ones. Second, the construction of a visual image incorporating the tactile data and the results of exploratory movements...*" (1956, p. 19). We are not certain that there is any essential difference between these two problems or processes but would suggest that there is a further process involved when identifying a shape-pattern in situations of this kind. This is a mediation process which involves the use of language and for present purposes may simply be called '*naming*.'

Seen in the extracts of reports of what Ss say in the Piaget and Inhelder experiments are the beginnings of attempts to schematize the experienced touch impressions into a meaningful form. S Ani is clearly trying to do this when she calls the triangular figure a 'planter' (a garden implement—the English word is probably better translated as 'dibber'). A child in the Page experiment describes the intertwined rings as 'two chickens.' This attempt to assign symbols or names to perceived objects and shapes raises important problems for experimental investigations of this kind. The acquisition of language may be considered as the increasing facility with, and range of, sets and sub-sets of words, some of which may be direct representations of objects in the spatial environment of the child and to relationships between them (Fisher, 1962). Hence, when presented with shapes, children may attempt to identify their tactile-impressions in terms of classification categories which are meaningful to them and to some of these categories they may be able to allocate a name as in the examples given above.

The studies of Piaget and Inhelder indicate that the frequency with which young children are able to recognise topological shapes is greater than that of linear shapes. It is necessary, however, to consider not only differences in the geometrical features of the shapes used, but also the possibility that an abstract shape may be classified as being one of a certain kind and that this classification, whether made overt or not, may influence subsequent classification of a comparison object. The two groups of shapes used appear to be such that the size of the ensemble of classification categories available to each is somewhat different. While only a very limited number of words appear to be available to Ss for classification of the linear shapes, a very much larger number appear to be available for classification of the topological shapes which have greater variety in their configurations.

In order to consider the hypothesis relevant to the influence of available classifying categories, sets of ten linear and ten topological shapes were prepared. Piaget and Inhelder, unfortunately, reproduce only rough sketches of some of the shapes used in their study. However, Page (1957) reproduces drawings of the shapes and photographs of some of them in the hands of children. Hence, by extrapolation, it was possible to estimate the approximate dimensions of the original materials. The shapes were cut from 3 mm. hardboard and were quite rigid and smooth to the exploratory touch of the palms and fingers. In order that there should be no difficulty in reproducing these shapes for future studies, sets of drawings of them, along with photographs of the objects and models used in Experiment I, have been published independently of the present paper (Fisher, 1963).

To each of the ten linear and ten topological shapes was allocated randomly a nonsense syllable taken from Glazes Lists (in Stevens, 1951, p. 542), of association value between 47 per cent. and 63 per cent. Twenty-six children, in the age range 1.7—5.0 years were used in the experimental group. Each S was shown a randomly selected set of three linear and three topological shapes and taught the nonsense 'name' of each shape to the criterion of being able to respond correctly twice by saying the name of the shape when it was indicated by *E* and by pointing to the shape when it was mentioned by name. During the course of this procedure Ss were not allowed to feel or touch any of the shapes. Each S was now placed in the situation used in Experiment I and asked to identify the shapes, which were placed into his hands in random order, either by saying its nonsense name or by pointing to drawings of the shapes mounted in the wooden frame. Each of the six shapes was presented on ten occasions, there thus being sixty trials for each S. Data were collected in the form of the number of correct responses either by naming, or pointing to, each shape. Nonsense syllables of a relatively high association value were chosen in order to ensure that the experiment would not be too time consuming for *E*, or too difficult for Ss. In order to minimise any effects due to differential association values, the syllables were reallocated to the shapes after the experiment had been completed on half the Ss in the experimental group.

A control group of twenty-six children in the age range 1.8—5.1 years was used also. The procedure used for control Ss was similar to that for the experimental group in that each child was allowed approximately the same amount of visual experience of the shapes but in the control situation he was asked to identify the shapes by pointing to them only.

TABLE 2
FREQUENCIES OF RECOGNITION OF LINEAR AND TOPOLOGICAL SHAPES.

EXPERIMENTAL GROUP					CONTROL GROUP						
S	Age		Frequency of recognition		Difference*	S	Age		Frequency of recognition		Difference*
	ys.	ms.	Linear	Topo-logical			ys.	ms.	Linear	Topo-logical	
1	1	7	12	2	- 1	10	1	8	7	16	+ 9
2	1	7	12	5	- 7	2	2	0	5	12	+ 7
3	1	10	22	10	-12	3	2	3	6	14	+ 8
4	1	11	24	11	-13	4	2	4	6	18	+12
5	2	3	18	8	-10	5	2	4	10	15	+ 5
6	2	3	13	5	- 8	6	2	4	14	12	- 2
7	2	5	21	14	- 7	7	2	5	8	14	+ 6
8	2	6	18	22	+ 4	8	2	7	9	15	+ 6
9	2	9	18	13	- 5	9	2	7	21	27	+ 6
10	2	9	24	21	- 3	10	2	9	17	17	=
11	3	1	23	28	+ 5	11	2	11	20	17	- 3
12	3	3	25	19	- 6	12	2	11	19	19	=
13	3	3	19	12	- 7	13	3	0	19	21	+ 2
14	3	3	30	22	- 8	14	3	0	16	22	+ 6
15	3	4	30	30	=	15	3	3	19	25	+ 6
16	3	6	27	23	- 4	16	3	3	17	12	- 5
17	3	7	21	7	-14	17	3	6	24	24	=
18	3	7	26	26	=	18	3	7	26	24	- 2
19	3	9	28	23	- 5	19	3	7	17	19	+ 2
20	4	0	30	30	=	20	4	1	29	28	- 1
21	4	3	26	27	+ 1	21	4	3	28	29	+ 1
22	4	5	26	29	+ 3	22	4	3	29	26	- 3
23	4	5	30	30	=	23	4	4	22	28	+ 6
24	4	9	30	30	=	24	4	7	24	24	=
25	5	0	30	29	+ 1	25	4	9	26	27	+ 1
26	5	0	30	30	=	26	5	1	29	28	- 1

$\bar{D}=4.04$
 $SE\bar{D}=1.05$
(One tailed test $V=25$).

$t'=3.85$
 $p'=0.001$

$\bar{D}=+2.54$
 $SE\bar{D}=0.85$
(One tailed test $V=25$).

$t'=2.99$
 $p'<0.01 >0.001$

LINEAR	
$\bar{X}_{\text{Exptl}} - \bar{X}_{\text{Contl}} = 5.63$	$t' = 2.95$
$SE\bar{X}_{\text{Exptl}} - \bar{X}_{\text{Contl}} = 1.91$	$p' < 0.001 > 0.002$
TOPOLOGICAL	
$\bar{X}_{\text{Exptl}} - \bar{X}_{\text{Contl}} = 1.04$	$t' = 0.48$
$SE\bar{X}_{\text{Exptl}} - \bar{X}_{\text{Contl}} = 2.16$	$p' > 0.10$
(One tailed tests $V=50$)	

* +ve topological > linear.
-ve linear > topological.

Again, in order to ensure that the overall pattern of recognition may be seen clearly, the data have been entered in Table 2 for individual Ss. These data appear to indicate fairly clear developmental trends in shape recognition ability. Firstly, all Ss, by the age of about 4 years, were able to recognise most of the shapes. Secondly, the preponderance of negative signs indicates that Ss in the experimental group were able to identify linear more readily than topological shapes. These data have been analysed using the technique applicable to consideration of differences between *correlated* means referred to in Experiment I. This indicates that the probability of the ensuing 't' value of 3.85 having occurred by chance is approximately 0.001. Thirdly, the preponderance of positive signs indicates that Ss in the control group, the original Piaget situation, were able to identify topological more readily than linear shapes. These data have been analysed using the technique referred to above also. Here the difference in performance is not as great as that in the experimental group but, nevertheless, the probability of the 't' value of 2.99 having arisen by chance is less than 0.01.

Although the only matching of Ss in the two groups was by age and this somewhat roughly, we may consider tentatively further differences in performance on the two shape groups. If the hypothesis that availability of a 'name' influences recognition ability is to be upheld, then it should also follow that frequency of recognition of linear shapes is *greater* in the experimental group than in the control group. Further, there should be *no* difference in frequencies of recognition of topological shapes between the two groups of Ss. In order to consider these two predictions, the data have been analysed further using a technique applicable to consideration of differences between *uncorrelated* means. In the case of linear recognition the probability of the emergent 't' value of 2.95 having occurred by chance falls between 0.01 and 0.02, a significant indication that the prediction concerning linear shape recognition is supported. In the case of topological recognition, while the difference in recognition is not statistically significant, it appears that the performance of Ss has been depressed slightly by teaching them the names of the shapes. Speculation as to the interpretation of this outcome is probably out of place with such data but it seems possible that the experimental procedure was such as to include operations which to some extent inhibited any implicit mediation processes involved.

The shapes presented to both experimental and control groups were identical. What was achieved by allocating the nonsense names to the shapes was that the sub-set of response categories was controlled in such a way that the probability of a response being available to each shape in the categories 'topological' and 'linear' was equal. While the differences between frequency of identification in the control group gives general support to the findings of Piaget and Inhelder, the differences in the experimental group is in a direction opposite to these. Hence, when the experiment is controlled in this way, it appears not only that the 'topological-primacy hypothesis' is false but that it requires to be replaced by one of 'linear-primacy.'

IV.—DISCUSSION.

Many criticisms have been levelled at the work of Piaget and his associates. The majority of these have been concerned with the non-representative nature of Ss in his experimental samples, with the lack of a satisfactory experimental design and with the almost complete disregard of the need for objective presentation of data. In spite of this, Piaget's contribution to our knowledge of the

nature of features of child development has been outstanding and, generally speaking, when his experiments have been repeated using more rigorous procedures, the findings have not been found to affect the overall argument greatly. It does, nevertheless, seem difficult to understand how Piaget came to suggest a theory of spatial development of the kind under consideration. He first (1955) described the development of the infant spatial world as being "... *The transition from chaos to cosmos* ..." (p. xiii) presumably indicating that the neonate's spatial world is completely undifferentiated or homogeneous. He argued later (1956), however, that "... *this (the developing) sensori-motor space is superimposed upon various pre-existing spaces such as the postural, etc.* ..." (p. 3). This is, of course, a logical contradiction since a pre-existing space of any kind implies non-randomness of some nature in, or feature of, the infant's spatial world and by definition a non-random state is not chaotic.

It seems extremely unlikely that human bipeds who achieve an upright posture at about the age of one year should lag behind by seven or eight years in the achievement of an operational spatial system, which is more or less linear. It is doubtful, however, if anyone has taken Piaget's suggestions as to teaching principles of topological geometry before linear or Euclidean geometry very seriously and the present findings indicate that even if this were practicable it would probably be unwise. In the course of conducting the present studies we became increasingly concerned about the fact that whatever processes are involved in the development of spatial operations, in situations of this kind at least, they appear to be complete by the age of about four years. The Ss used in the present experiments were probably, however, of slightly higher than that of a randomly selected group and, although they were not given intelligence tests, a subjective assessment of their average intellectual level would be about 110 I.Q. points. Experiment II has now been repeated on a sample of children attending an educationally subnormal school and thus of I.Q. somewhat less than 70. While the pattern of results is the same as that for the original sample, it is noteworthy that equal frequency of recognition of 'topological' and 'linear' shapes does not appear until between 10 and 11 years. Hence intelligence does appear to influence performance in this situation.

Further experiments are now under way concerned with children's perception of phenomenal causality, illusions and the spatial verticals. Certain features of these appear to be critical in considering the developmental theories of Piaget and others.

V.—REFERENCES.

- FISHER, G. H. (1962). Phenomenal causality in conditions of intrasensory and intersensory stimulation. *Amer. J. Psychol.*, 75, 321-323.
- FISHER, G. H. (1963). *Objects, Models and Two-Dimensional Shapes for Use in Visual and Tactile-Kinaesthetic Shape Perception Experiments*. Newcastle: University of Durham, Institute of Education.
- PAGE, E. I. (1957). *Haptic Perception*. University of Birmingham, Diploma in Education Dissertation (unpublished).
- PAGE, E. I. (1959). *Haptic Perception: A consideration of one of the investigations of Piaget and Inhelder*. *Educ. Rev.*, 11, 115-124.
- PIAGET, J. (1955). *The Child's Construction of Reality*. London: Routledge and Kegan Paul.
- PIAGET, J., and INHELDER, B. (1956). *The Child's Conception of Space*. London: Routledge and Kegan Paul.
- PIAGET, J. (1962). Comments. M.I.T.
- REVESZ, G. (1950). *Psychology and Art of the Blind*. London: Longmans.
- STEVENS, S. S. (1951). *Handbook of Experimental Psychology*. New York: John Wiley.

(Manuscript received 1st June, 1964.)

SEX DIFFERENCES IN ATTITUDES AND ATTAINMENT IN JUNIOR SCHOOLS

BY MILES WISENTHAL

(Institute of Education, McGill University)

SUMMARY. An investigation was carried out among 1,164 boys and 1,085 girls in single-sex and mixed schools. Significant differences in favour of the girls were found in I.Q. attainment, and in attitude to school. It was also noted that attitudes to school grow more unfavourable from first to fourth form. Where there is a very substantial deterioration in attitude there is a marked lack of improvement in attainment relative to I.Q.

I.—INTRODUCTION.

A SUBSTANTIAL number of studies have been carried out which demonstrate the academic superiority of girls over boys in the first six to eight years of school. In Great Britain the following investigators, among others, have shown that girls achieve better results than boys in most subjects of the junior school: Emmett (1954); Clark (1955); Vernon, O'Gorman, and McLelland (1955); Richardson (1956); Blandford (1957); Yates and Pidgeon (1957); Pidgeon (1960); and France (1964). In the United States the published research is very extensive indeed. Goodenough (1954), after a thorough survey of the findings on sex differences in school performance, concluded that on the average, girls achieve better results than boys.

The numbers involved in the studies cited above range from a few hundred to several thousand. For truly large numbers one must examine the school census data that are available. The United States Census (1962), the report of the Dominion Bureau of Statistics of Canada (1960), and the New Zealand (1959) school census information all show with striking clarity the extent of the difference in school attainment between boys and girls. The facts are made plain when the age-grade tables which are provided in these publications are analysed. There are highly significant differences in favour of boys in the number of children shown as over-age for all school grades above three (9-year-olds). Conversely, there are highly significant differences in favour of girls in the numbers shown for under-age or at-age for the same grades. The principal reason for children being over-age is that grades are repeated by those who fail to achieve a passing standard of performance. In the United States, New Zealand and Canada, this occurs with much greater frequency among boys than it does among girls. That this phenomenon is universal is shown by Wall, Schonell, and Olson (1962) in their study on school failure for UNESCO.

The most common reason advanced to explain the disparity in performance between boys and girls is the more rapid maturation of girls. Other reasons advanced include such possible explanations as that of Goodenough (1954) in which she refers to the greater docility of girls; Richardson (1956) suggests that boys are more intolerant than girls of uninteresting lessons; Carter (1952) proposed that all teachers show a marked bias in favour of girls when assigning grades. It is not the purpose of this paper to quarrel with the hypotheses advanced by others to explain sex differences in school performance, but rather to suggest that there may be factors affecting school attainment which, hitherto, have been overlooked.

The work of Gardner and Lambert (1959) has demonstrated that there can be a powerful link between school attainment in a subject and the attitude of the learner. If there is a causal relationship between attitude and the learning of a single subject, it would seem reasonable to expect that attitudes to school would effect total school performance.

In the light of what is already known about sex differences in junior school (ages 7-11) attainment, and the clear-cut findings of Gardner and Lambert on the effect of attitude to second-language acquisition, it was decided to examine what relationship, if any, exists between total school performance and attitudes. The selection of junior school children as the subjects for study is based on the belief that the attitudes of secondary school pupils arise out of their early school experiences. Thus, any long-range plans for changes in secondary schooling must begin with a complete understanding of the role of primary education in attitude formation.

In view of the fact that differences in attitude might be related to whether the school was an all-boys', all-girls', or mixed, it was decided to include in the investigation a proportion of each type. Through the co-operation of the Kent Education Committee, three boys', three girls', and six mixed junior schools were made available to the writer. These twelve schools produced a total of 1,164 boys and 1,085 girls, apportioned as shown in Table 1. It will be noted that no third forms are included: the reason for the omission is an administrative one related to a larger study of which the present report makes up one part.

TABLE 1

NUMBER OF PUPILS BY SEX AND BY FORM IN THREE BOYS', THREE GIRLS', AND SIX MIXED SCHOOLS.

	Boys' Schools	Girls' Schools	Mixed Schools	
Form			Boys	Girls
1	196	205	179	149
2	214	196	178	166
4	214	204	183	165
Total	624	605	540	480

The measures of attainment used in the junior schools of Kent are well designed standardised tests for which new norms are prepared annually. The intelligence tests are of similar high quality. In forms one, two, and three, the testing is carried out in the second half of the Summer term; in the fourth form testing is completed by the end of January. All tests, both attainment and intelligence, have a mean of 100, a standard deviation of 15, a floor of 65, and a ceiling of 135. In the first form the I.Q. is based on a non-verbal test of intelligence; in the other form it is the mean of a verbal and non-verbal intelligence test. The attainment measure for forms one and four is the mean of a reading comprehension and a reasoning arithmetic test; in the second form it is the mean of a reading comprehension, spelling, reasoning arithmetic and mechanical arithmetic test.*

* For a complete account of the tests and testing procedures, see N. France in this *Journal*, 34, 19-33, 1964.

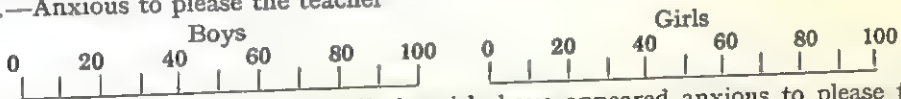
The measurement of attitudes is a difficult procedure even under the most favourable circumstances. In view of the difficulties likely to be encountered with young children, it was decided to ask the teachers to complete the questionnaire which follows :

ATTITUDINAL DIFFERENCES BETWEEN BOYS AND GIRLS IN THE JUNIOR SCHOOL.

As we are all aware, the attitudes that children bring to school exert some considerable influence on their social behaviour and on their willingness to learn. A study is being undertaken to see if there are measureable differences in the attitudes of boys and girls to schools and teachers. It is believed that more realistic appraisal of attitudes can be made by teachers who have an intimate knowledge of children's school behaviour, than by the girls and boys themselves.

To complete the attached form objectively, it is necessary that you consider your total teaching experience. Think of a group of 100 boys and 100 girls, and estimate the number of each that would show the characteristic stated. For example, item number one appears as follows :

1.—Anxious to please the teacher



If, in your experience, nearly all the girls have appeared anxious to please the teacher and only a small number of boys have shown the same trait, then you should place a check mark at 90, or 100 for girls, and a check mark at 10, or 20 for boys. Please make sure that each item is checked twice, once for boys and once for girls.*

Your name

Years of teaching in junior school..... Infant school.....

Other.....

Number of boys in class

number of girls.....

Your co-operation is very much appreciated.

The twenty items used in the questionnaire were :

- | | |
|-----------------------------------|-------------------------------------|
| 1.—Anxious to please the teacher. | 11.—Appear to dislike the teacher. |
| 2.—Annoy the teacher. | 12.—Disrupting of class routines. |
| 3.—Lazy about doing school work. | 13.—Careless in work habits. |
| 4.—Dislike school. | 14.—Appear anxious to learn. |
| 5.—Inattentive to lessons. | 15.—Indifferent to teacher. |
| 6.—Think teacher unfriendly. | 16.—Polite. |
| 7.—Take pride in school work. | 17.—Rude. |
| 8.—Stubborn. | 18.—Disobedient. |
| 9.—Like to help the teacher. | 19.—Co-operative in class projects. |
| 10.—Will not persevere. | 20.—Considerate of teacher. |

It might be quite reasonable to argue that the results obtained from the above questionnaire are a more valid measure of teachers' attitudes than of children's ; there are two reasons why this makes no real difference in the outcome of the investigation. Firstly, it seems very likely that primary children's attitudes to school are, in a large measure, a result of the attitudes of the teachers to children ; thus, one should be able to use the obtained results as a reasonably accurate assessment of children's attitudes to school. Secondly, if one chooses not to accept this argument, then by merely changing the title of this investigation to " Sex Differences in Teachers' Attitudes and Pupil Attainment," the writer will not take offence.

* Teachers in single-sex schools were asked to rate either boys or girls only.

In order to simplify the scoring of the questionnaire all the items were scored negatively. Thus, a respondent who stated that 90 per cent. of the boys are "anxious to please the teacher," was awarded a score of 10 for the item because this response can be taken to mean that 10 per cent. of the boys are not anxious to please the teacher. If a respondent marked the scale at 10 per cent. for a negative item such as 'dislike school,' this was taken at face value and scored as 10. The sum of all the item scores represents the degree of negativity of the pupils, either boys or girls. Theoretically, a teacher could mark the questionnaire so that a zero score would result, indicating a complete absence of any negative feelings or a highly favourable attitude; conversely, it would be possible to achieve a score of 2,000, an extremely negative score indeed. In actual fact, the scores ranged from a low (favourable) of 40, to a high (unfavourable) of 890. The median for boys was 407 and for girls 249. Not marked differences were shown in the questionnaires completed by male or female teachers, or between those completed by married and unmarried teachers.

Table 2 compares the total sample of boys and girls for I.Q. and attainment. It will be noted that girls show both higher I.Qs. and attainment. What is of particular interest is the fact that the mean attainment for boys is slightly below their mean I.Q., while that of the girls is 1.62 above their mean I.Q.

TABLE 2
MEAN I.Q. AND ATTAINMENT FOR BOYS AND GIRLS IN FORMS 1, 2, AND 4.*

	Boys			Girls			Difference B—G
	N	Mean	s	N	Mean	s	
I.Q. . .	1162	100.66	13.05	1085	102.75	13.25	-2.09 > .001
Attain.	1169	100.47	14.00	1089	104.47	13.35	-4.00 > .001

Attitude questionnaires were completed for boys by forty-four teachers and forty-nine teachers completed those of the girls. The magnitude of the difference in means between the attitudes reported for boys and for girls is shown in Table 3. The reader is reminded that the attitude scores, like golf scores, are at their worst when high. While this highly significant difference in favour of the girls is of no little interest, what is perhaps of greater importance is the differences between the various types of school, and the changes in attitude that occur from the end of the first-year in junior school to the fourth year. Table 4 contains this information.

TABLE 3
MEAN ATTITUDE SCORES FOR BOYS AND GIRLS IN FORMS 1, 2, AND 4.

Boys			Girls			Difference B—G
N**	Mean	s	N**	Mean	s	
44	412.50	166.50	49	245.40	154.00	+167.10 < .001

* The difference in the N's between I.Q. and attainment results from a small number of children for whom test scores were not available.

** The N used is that of the number of teachers who completed the questionnaire.

TABLE 4

DIFFERENCES BETWEEN FIRST FORM AND FOURTH FORM IN MEAN I.Q., ATTAINMENT, AND ATTITUDE IN BOYS', GIRLS', AND MIXED SCHOOLS.

	First Form				Fourth Form			
	N	I.Q.	Attain.	Attit.	N	I.Q.	Attain.	Attit.
Boys' Schools...	196	101.49	100.49	362	214	98.75	100.37	434
Girls' Schools ..	205	103.29	106.73	104	204	104.18	105.79	252
Mixed Schools :								
Boys	179	101.40	100.92	421	183	99.32	100.94	484
Girls	149	101.70	102.93	292	165	98.40	101.45	339

It will be noted that the first forms in the girls' schools show the highest attainment of all the first forms; they also show the greatest excess of attainment over I.Q. Their mean attainment exceeds their mean I.Q. by 3.44. It is also noteworthy that their attitudes are reported as being considerably more favourable than those of the other first forms. The first forms in the boys' schools do not enjoy the same degree of academic success; their mean attainment is 1.00 below their mean I.Q., and they show substantially more negative attitudes to school than do the first form girls.

The boys and girls in the mixed schools have practically identical mean I.Qs. in the first form, but whereas the boys' mean attainment is .48 below their mean I.Q., the girls show a mean attainment which is 1.23 above their mean I.Q. There is, too, a considerable difference in attitude in favour of the girls.

No doubt it will have been noted that the mean I.Q. in the girls' schools is significantly higher at both the first and fourth forms, than those of any of the other girls or boys. This is difficult to explain inasmuch as the boys and girls at the single-sex schools all come from the same neighbourhoods, and the intelligence tests have been vetted for sex bias. Is it possible that more favourable attitudes to school are influencing intelligence test performance?

The girls' schools retain their lead into the fourth forms. At this point they have higher mean I.Qs., higher attainment means, and more favourable attitudes than any of the other fourth forms. However, when the difference in attitude between the first and fourth forms in girls' schools is compared, it will be seen that their mean attitude score has increased substantially, that is, worsened. The mean attitude score of 104, which they show in the first form, places these girls at the 26th percentile for all girls; in the fourth form their mean of 252 places them at the 52nd percentile. Further, it will be seen that the mean attainment in the fourth form exceeds the mean I.Q. by only 1.61 as compared to 3.44 in the first form.

In the boys' schools the mean attainment related to mean I.Q. moves from -1.00 in form one to +1.62 in form four. Their mean attitude score moves from the 44th percentile to the 59th, from the first to the fourth forms. In this case, a relatively narrow movement in the attitude score is accompanied by a substantial improvement in attainment relative to I.Q. Boys in mixed schools show nearly an identical pattern. Mean attainment relative to I.Q. moves from -48 to +1.62, and their mean attitude score moves from the 55th to the 67th percentile.

Girls in mixed schools show the smallest change in attitude and also show the highest attainment relative to I.Q. for all the fourth forms. Their mean attainment moves from +1.23 to +3.05, and their mean attitude score move from the 62nd to the 68th percentile.

In the three cases in which the attitudes to school do not show any really gross change one observes a substantial improvement in attainment relative to I.Q., from the first to the fourth form. In the girls' schools, however, where there is a marked deterioration in attitude, one notes a marked drop in attainment relative to I.Q.

Inasmuch as there are substantial differences in attitude between each of the first forms, it might be useful to think of each group as having its own established attitude norm against which changes can be measured. Each of these groups has created a basic attitude level as a result of experiences which are specific to its sex and type of school, thus inter-group comparisons may be quite invalid. What appears to have some validity is the magnitude of the change in attitude within the group over the four-year period of junior school.

The relationship between attitude and attainment is difficult to explain. A possible explanation that suggests itself is that children's attitudes to school are formed on their initial contact with their first form teachers; these attitudes establish a set of expectancies. As long as these expectancies remain more or less constant, children's school performance will improve with maturity. On the other hand, if children begin in the first form with a highly favourable but unrealistic set of attitudes, this will lead to the establishment of expectancies which cannot be confirmed or realized. The disillusionment which accompanies the failure to confirm expectancies leads to a reduction in performance as was noted in the girls' schools. If one observed a similar pattern among the girls in mixed schools, it would be easy to explain this phenomenon in terms of the onset of puberty, and the concern which girls at this age show in social rather than academic achievement. But it appears from this limited investigation to occur only in the girls' schools, which suggests that the first explanation may be correct.

The relationship of the attitudes of teachers to pupils and pupils to school is a matter of considerable importance to the educational enterprise. The complexity of the problem makes it difficult to provide clear-cut answers to the questions that present themselves, but this should not deter further research in this important area.

II.—REFERENCES.

- BLANDFORD, J. S. (1957). *Standardised Tests in Junior Schools with Special Reference to the Effects of Streaming on the Constancy of Results*. Unpublished M.A. Thesis, University of London.
- CLARK, HILDA M. (1955). *The Effect of a Candidate's being Above or Below the Median Age Upon His Chances of Selection for the Grammar School in the Eleven Plus Selection*. Unpublished M.A. Thesis, University of London.
- DOMINION BUREAU OF STATISTICS (1960). *Student Progress Through the Schools*. Ottawa: The Queen's Printer and Controller of Stationery.
- EMMETT, W. G. (1954). Secondary modern and grammar school performance predicted by tests given in primary schools. *Brit. J. Educ. Psychol.*, 24 (2), 91-98.
- FRANCE, N. (1964). The use of group tests of ability and attainment: a follow-up study from primary to secondary school. *Brit. J. Educ. Psychol.*, 34 (1) 19-33.
- GARDNER, R. C., and LAMBERT, W. E. (1959). Motivational variables in second language acquisition. *Cdn. J. Psychol.*, 13, 266-272.

- GOODENOUGH, FLORENCE L. (1954). The measurement of mental growth in childhood. In the *Manual of Child Psychology*. Ed. L. Carmichael., 482-483. New York: John Wiley and Sons, Inc.
- NEW ZEALAND (1959). *Age and Classification of Pupils at Public Primary Schools*. Report of the Minister of Education for the year ended 31st Dec., 1958. Wellington: R. E. Owen, Government Printer.
- PIDGEON, D. A. (1960). A national survey of the ability and attainment of children at three age levels. *Brit. J. Educ. Psychol.*, 9, 13-31.
- RICHARDSON, S. C. (1956). Some evidence relating to the validity of selection for grammar schools. *Brit. J. Educ. Psychol.*, 26, 15-24.
- UNITED STATES CENSUS (1962). *1960 Census of Population: Supplementary Reports*. School enrolment of the population of the United States: 1960. PC (S1)—29 Washington: U.S. Department of Commerce.
- WALL, W. D., SCHONELL, F. J., and OLSON, W. C. (1962). *Failure in School: An International Study*. Hamburg: UNESCO.
- YATES, A., and PIDGEON, D. A., 1957. *Admission to Grammar Schools*. London Newnes and Co.

(Manuscript received 16th June, 1964.)

RESEARCH NOTES

MULTIPLE REGRESSION EQUATIONS FOR PREDICTING READING AGE FROM CHRONOLOGICAL AGE AND WISC VERBAL I.Q.

By FAY FRANSELLA AND DAVID GERVER*

(*Department of Psychology, Institute of Psychiatry, University of London*)
(*Maudsley Hospital, Denmark Hill, London, S.E.5*)

SUMMARY. To assist clinical and educational psychologists in comparing a child's reading attainment on the Schonell Graded Word Reading Test (GWRT) with that expected from one of his age and I.Q., three multiple regression equations are presented. These enable the psychologist to predict Reading Age on Schonell's GWRT from a given chronological age and WISC Verbal I.Q.

The equations were calculated from data relating to 500 children tested at the Children's Department of the Maudsley Hospital, with the Schonell Graded Word Reading Test and the Wechsler Intelligence Scale for Children. Separate predictor equations are presented for age groups 6.2-9 years, 10-12 years, 13-15 years, because of the differing correlations between reading age and I.Q., which are found with increase in chronological age.

INTRODUCTION.

The educational or clinical psychologist may sometimes wish to compare a child's intelligence test score with his reading age and will only be able to make such statements as "the child's reading age appears to be consistent with his I.Q." Such statements are usually based on rule-of-thumb methods rather than on any statistical data. In the case that prompted this study the psychologist was asked by the Consultant Psychiatrist whether a reading age of 12 years (on the Schonell Graded Word Reading Test, 1950) for a girl of 16 with a WISC I.Q. of 60 was higher than one might expect for someone of this age and intellectual level. Since there were no data available whereby such questions could be answered, multiple regression equations were computed for predicting reading age (derived from Schonell's G.W.R.T.) from chronological age and W.I.S.C. Verbal Scale I.Q. This predicted reading age could then be compared with that obtained and the significance of the difference computed should this be considered necessary.

Ravenette (1961) discusses a similar approach to assessment of reading retardation and presents data for predicting reading age on the Schonell G.W.R.T. from scores on the Crichton Vocabulary Scale.

SAMPLE DATA.

The records of 500 boys and girls within the age range of 6.2 years to 15 years were alphabetically extracted from the files. These children had been tested in the Children's Department of the Maudsley Hospital between 1952 and 1962 with the W.I.S.C. and Schonell's G.W.R.T. All scores on the G.W.R.T. were corrected as per the Ministry of Education pamphlet No. 18 entitled "Reading Ability."

In view of the general observation that the correlation between reading age and I.Q. increases with increase in chronological age, the total sample in this study was divided into three chronological age groups, details of which are given in Table 1. Separate calculations for males and females were not made. The differences between mean Full Scale I.Q. scores for the 6.2-9 year and the 13-15 year age groups is significant ($p < .05$), as is the difference between mean Verbal I.Q. scores for these age groups.

* Now with the M.R.C. Psycholinguistics Research Unit, Institute of Experimental Psychology, Oxford.

TABLE 1

MEANS AND STANDARD DEVIATIONS FOR TOTAL SAMPLE AND THREE AGE SUB-GROUPS

	AGE RANGE							
	6·2—9		10—12		13—15		6·2—15	
	N=118		N=224		N=158		N=500	
	\bar{x}	S.D.	\bar{x}	S.D.	\bar{x}	S.D.	\bar{x}	S.D.
Chronological Age	8·72	0·82	11·56	0·91	13·96	0·63	11·65	2·09
Reading Age ..	8·01	2·24	10·75	3·18	12·57	3·10	10·66	3·38
Verbal Scale I.Q.	98·36	14·46	95·83	18·96	93·85	16·47	95·80	17·26
Full Scale I.Q. ..	97·85	15·68	94·85	19·95	93·76	18·30	95·20	18·53

THE EQUATIONS.

In Table 2 are presented three multiple regression equations for prediction of Schonell G.W.R.T. Reading Age for the three age groups.

TABLE 2

Age	Multiple Regression Equations	Standard Error of Estimate
6·2—9	$-8·44 + (0·98 \times \underline{CA}) + (0·085 \times \underline{I.Q.})$	1·90
10—12	$-7·68 + (0·64 \times \underline{CA}) + (0·117 \times \underline{I.Q.})$	2·24
13—15	$-10·86 + (0·72 \times \underline{CA}) + (0·114 \times \underline{I.Q.})$	1·96
Three multiple regression equations for predicting reading age on the GWRT for three age groups.		

The Verbal Scale I.Q. was used with chronological age as a predictor variable rather than the Full Scale I.Q. since (as will be seen from inspection of Table 3) there is a slight gain in the accuracy of prediction due to lower errors of estimated reading age (S.E.est) and a higher proportion of the total variance is accounted for ($r^2_{1.23}$).

TABLE 3

Age	Full Scale		Verbal Scale	
	S.E.est	$r^2_{1.23}$	S.E.est	$r^2_{1.23}$
6·2—9	1·79	0·37	1·79	0·37
10—12	2·38	0·44	2·24	0·50
13—15	2·11	0·54	1·96	0·60

Standard errors of estimate and variance accounted for when the Verbal Scale I.Q. or Full Scale I.Q. is used with chronological age as predictor variables.

As an example of the use of the equation consider the case of a child aged 14.2 years with a WISC Verbal I.Q. of 80 and an obtained Schonell Reading Age of 10.1 years. Is this obtained Reading Age to be expected in a child of this age and I.Q.?

$$\begin{aligned}\text{Predicted Reading Age} &= -10.86 + (0.72 \times \text{CA}) + (0.114 \times \text{I.Q.}) \\ &= -10.86 + (0.72 \times 14.2) + (0.114 \times 80) \\ &= -10.86 + 10.22 + 9.12 \\ &= 8.48\end{aligned}$$

Test for abnormality of the difference between obtained and expected reading age as follows:

$$\begin{aligned}\frac{\text{Obtained RA} - \text{Predicted RA}}{\text{S.E.}_{\text{est}}} &= \frac{10.10 - 8.48}{1.96} \\ &= \frac{1.62}{1.96} = .827 \\ &= 0.827 \quad P > .05\end{aligned}$$

One might thus reasonably conclude that a Reading Age of 10.1 years is to be expected in a child of 14.2 years of age with a verbal I.Q. of 80.

DISCUSSION.

The use of separate equations for each age group will be seen to be justified on inspection of Table 4 in which are compared the partial correlations of reading age with chronological age, I.Q. held constant and of reading age with I.Q., chronological age held constant.

TABLE 4

Age	Full Scale I.Q.		Verbal Scale I.Q.	
	$r_{\text{RA} \times \text{CA. I.Q.}}$	$r_{\text{RA} \times \text{I.Q. CA}}$	$r_{\text{RA} \times \text{CA. I.Q.}}$	$r_{\text{RA} \times \text{I.Q. CA}}$
6.2—9	0.43	0.54	0.41	0.55
10—12	0.25	0.65	0.25	0.70
13—15	0.25	0.72	0.22	0.76
The Partial Correlations for both Full Scale and Verbal Scale I.Q.				

It can be seen that the correlation of reading age with I.Q. increases with increase in chronological age, while the correlation of reading age with chronological age decreases with increase in age. The fact that the present sample is drawn from an abnormal population, means that no generalization about this finding can be made. However, the mean Verbal Scale I.Q. of 95.8 and standard deviation of 17.26 are consistent with those found in normal populations. Netley, Rachman and Turner (in press) found the reading age, as measured on the Neale reading test, to be just under a year below that of the chronological age in a sample of children from one primary school run by a local authority. The findings that the difference between the Verbal Scale I.Q.s. for the first and last age group is significant, and that the $\text{RA} \times \text{I.Q.}$ correlation increases with age, are enough to justify the using of different equations and to militate against the use of these equations with other than a clinic population.

A word about the interpretation of the difference between obtained and predicted reading ages. Should a test of the significance of the difference be required, a one tail test is appropriate since the direction of the difference will be evident from the data. It must be borne in mind that these equations are based on data derived from a possibly abnormal sample of children so that one might be, in effect, testing one abnormal score against another. For this reason the psychologist might consider the .05 level of significance too severe a test. From a practical point of view, the greatest use for these equations is that they allow a fairly accurate assessment as to what the reading age of a child should be in relation to his age and I.Q.

ACKNOWLEDGEMENTS.—The authors wish to acknowledge their indebtedness to Mr. H. Gwynne Jones and Dr. A. E. Maxwell for their advice in the preparation of this study. They also wish to acknowledge the assistance of the Computer Centre, The University of California, Berkeley, California, where the computations were carried out by one of the authors (D.G.) during the year 1962-63.

REFERENCES

- EXEKIEL, M. (1941). *Methods of Correlation Analysis*. New York: John Wiley and Sons.
- MCNEMAR, Q. (1962). *Psychological Statistics*. New York: John Wiley and Sons.
- NETLEY, C., RACHMAN, S., and TURNER, R. K. (1965). *The Effect of Practice on Performance in a Reading Attainment Test*. (In press.)
- RAVENETTE, A. T. (1961). An empirical approach to the assessment of reading retardation, vocabulary level and reading attainment. *Brit. J. Educ. Psychol.*, 31, 96-103.
- SCHONELL, F. J. (1950). *Diagnostic and Attainment Testing*. Edinburgh: Oliver and Boyd.

(Manuscript received 24th June, 1964.)

A COMPARISON OF THE CONCEPTS OF CATTELL AND EYSENCK

By C. J. ADCOCK

(Victoria University of Wellington, New Zealand)

I.—INTRODUCTION.

FACTORIAL studies of personality have been many and varied but in the U.S.A. no one has made more contributions than R. B. Cattell. In Britain, the name most associated with this work is that of Eysenck. Both men are well known beyond their own countries and both at times seem to have been talking about the same things. But one easily gets the impression that, when they do talk about the same things, what they say is different and that when what they say is the same, it is about different things. In the interests of research it is worth considering how they differ and why. No attempt will be made to cover all the work of either man. Attention will be restricted to personality concepts based on factorial and related studies.

This is, perhaps, somewhat unfair to Eysenck whose major contention is that factor analysis is not enough. For him the stress is upon the hypothetico-deductive method in which factor analysis plays a part but requires to be integrated with other studies of a more causal nature. His more recent work has been almost entirely in this wider sphere. While Cattell has been concerned to introduce greater precision into the factorial approach, Eysenck has preferred to use it only as an introductory tool whose products then require shaping with finer instruments. Much of his current work is concerned with this type of development in connection with his introversion factor. We can be concerned with this only in so far as it confirms one or other factorial interpretation. To this extent we may, therefore, do him an injustice and we would ask the reader to bear in mind that a balanced picture requires consideration of the non-factorial aspects also.

Our own sympathy is very much with Eysenck's stress on the wider approach, but we would prefer to see this linked more intimately with factorial methods at all stages. We see indication of a development in this direction in the recent paper in which Eysenck collaborated with his wife in a factor-analytic paper designed to clear up a problem relating to divergent concepts of introversion (Eysenck and Eysenck, 1963), and in other recent papers (Eysenck and Eysenck, 1962a, 1962b).

Basic to the disagreement between the two men in their factor analytic studies is a methodological one relating to the number of factors to be extracted. Eysenck's major contributions in this area were made during the early post-war years when most British factor-analysts were still very suspicious of Thurstone's notions of rotating to simple structure. Centroid factors account for decreasing amount of variance as further factors are taken out so rotation dictates the extraction of few factors. In his original research, Eysenck (1944) played safe and concentrated on two factors although he took out four. These were general neuroticism and hysteria-dysthymia and accounted for 14 per cent. and 12 per cent. of the variance, respectively. By contrast, Cattell (1946) took out twelve factors accounting for 40 per cent. of the variance. Later research adds four more factors.

What is at stake here is not a simple matter of how many factors should be extracted: it is rather the matter of rotation. If we are to interpret centroid factors, the ones we omit to extract will almost certainly be of less importance than the ones we have extracted, but if we are dealing with rotated simple-structure factors, stopping the factor analysis too soon will distort the rotated factors. Cattell must take out all significant factors. Eysenck may stop at any stage he chooses.

The study of plasmodes such as the box data of Thomson (1946) shows that it is the oblique rotated factors which correspond to the length, breadth and depth of the physical model, while a second order factor arising from the inter-correlation of these factors corresponds to the restriction in sampling which results from measuring only boxes which have been actually built for utilitarian purposes. Many theoretically possible boxes remain unconstructed in practice and this introduces a general size factor.

This situation of first-order factors which are themselves inter-related in some way is obviously a frequent one and one which will be adequately revealed only by oblique rotation which, be it noted, will indicate higher-order factors only if they are potentially present. If the box data are obtained from purely theoretical boxes the size factor disappears. In the light of this we would argue that Cattell's analysis into twelve, and later sixteen, oblique factors with four second-order factors is most likely to have psychological significance. It permits but does not require correlated primary factors.

Does it follow that Eysenck's two centroid factors will be completely misleading? This is not necessarily so at all. Insofar as the centroid factor takes out at each stage, what is most general to the matrix we should find that the first few centroid factors are likely to be influenced largely by the second-order factors. This is, in fact, the case. The two centroid factors of Eysenck correspond to the two largest second-order factors of Cattell. The latter calls them anxiety and extraversion or *exvia*. Eysenck elsewhere calls his second factor extraversion while the relation between anxiety and neuroticism seems obvious enough.

It would appear then that there are broad lines of agreement. Eysenck's procedure gives a less precise measure of the second-order factors. The lack of precision is undesirable but should not, in itself, lead to grave differences in the interpretation of the factors. The extent of these differences has now to be looked into.

II.—ANXIETY AND NEUROTICISM.

Eysenck called his first factor neuroticism but he would appear to regard its main constituent as being an emotional sensitivity which makes its possessor more vulnerable to the stress of living. Cattell (Cattell, *et al.*, 1962) is inclined to argue that neuroticism relates to eight out of sixteen of his 16 P.F. first-order factors and two second order factors. This looks like disagreement on the grand scale, but it may be more obvious than real. It is quite understandable that had Eysenck continued his factorization for another fourteen factors he might have found some further factors related to neuroticism. His deliberate policy ensured that only broad second-order factors could be measured. As already suggested, his *neuroticism* should, therefore, correspond to Cattell's *anxiety*. Both statistical evidence and the factor descriptions offered by the two

writers seem to indicate that this is, in fact, so. Within the restrictions of the differences in factorial approaches there appears to be full agreement in the evidence so far.

What, then, is disagreed? Is it just a matter of words? Eysenck uses the term 'neuroticism' to describe the factor because generally the items which load highly on the factor are those which best distinguish neurotics from normals. But Cattell too finds strong evidence for a similar relationship. All that he is arguing is that *there are other factors involved also*. Eysenck would hardly be disposed to quarrel with this. He would merely query the importance of the other factors. At the most he might argue that the extraversion-introversion factor should not be regarded as one of the factors related to neuroticism, but even here the superficial evidence may be somewhat misleading since his centroid factors are necessarily orthogonal while Cattell's second-order-factors are in some degree correlated.

So far our study has shown that Cattell's research supplements rather than opposes that of Eysenck. It may further be shown that the neurotic variance contributed by Cattell's eight first order factors is somewhat less than indicated by his data. He provides (1962, p. 43) sten means based on 201 subjects for the 16 P.F. factors. The three largest deviations from normal subjects are +2.8 for guilt proneness, +2.4 for ergic tension and -2.4 for ego-strength, but these are the three major contributors to the second-order anxiety factor which itself shows a deviation of +2.1. There is nothing surprising about the fact that the sten deviation of the major second order factor is equal to the average deviation of the factor scores on which it is based; but it does emphasise that this evidence leaves wide open the question as to what is the amount of variance contributed directly by the second-order factor of anxiety and how much is contributed by the individual first-order factors. It might well be that anxiety as a second-order factor is just a measure of emotional reactivity which largely determines the degree of guilt-proneness, ergic tension and ego strength and which is itself the greatest direct contributor to neurosis. Eysenck would doubtless go all the way with such a suggestion; Cattell would regard it as an exaggeration but only further work of his own kind will decide what degree of exaggeration may be involved.

But we have still to look at the 16 P.F. factors which are not involved in the second-order anxiety factor. Again listing in order of sten deviations from normals we have surgency -1.9, protension +1.7, dominance -1.6, shyness -1.2, and tender-mindedness +1.2. All these should be significant. We are struck at once by the possibility that some of these may relate to effects rather than causes of neuroticism. Surgency in its negative form involves depression which may often be a neurotic product. Shyness and protension (suspiciousness) could operate either as cause or effect. Lack of dominance and tender-mindedness seem more obviously causal factors.

This raises an important point. If we speak of the predictive capacity of a test of neuroticism we may be referring to the capacity of the test to predict what would be the outcome of psychiatric diagnosis now or we may be referring to the capacity of the test to predict which of our testees will be likely to develop neuroses in the future. These are entirely different things and must not be confused. By analogy a count of skin spots might be an objective index of measles in the first sense but quite useless for predicting measles in the second sense. Keeping this point firmly in mind would obviate a number of confusions. Both Cattell and Eysenck seem usually concerned with prediction in form one.

III.—OBJECTIVE TESTS.

So far we have been concerned with evidence from questionnaire tests. Both Eysenck and Cattell have sought evidence from objective tests. The former has found several objective tests which differentiate neurotics from normals (1950, 1957) and Cattell (1962) has found that similar tests load a factor which distinguishes neurotics. Perhaps we have agreement for once. To clinch the matter, Cattell threw in six psychiatric ratings of anxiety but they turned out to have insignificant relationship! Nevertheless, Cattell was himself sufficiently impressed to accept this factor as the major neurotic contributor in this series of 21 objective factors. This implies that he would consider U.I. 23, as it becomes designated in his system, to be on the same level as the second-order anxiety factor which corresponds to U.I. 24. The real disagreement here then is simply in the use made of words. Eysenck uses the title *neuroticism* for what he finds to be the broadest contributor to the actual condition of neurosis while Cattell's corresponding objective factor is only one of which characterise neuroticism. And yet we may be doing Cattell an injustice here. Strictly, his U.I. 23 factor is labelled 'High mobilization *versus* neurotic regressive debility' or more briefly, 'neural reserves.' As such it should be generally acceptable. If Eysenck would keep to the less controversial name for his general neuroticism factor (e.g., emotional reactivity) the position would be less confusing to the general reader.

We have now to take note of a further consequence of Cattell's more exhaustive system of analysis. His objective tests yield even more factors than did his questionnaire tests and again some 50 per cent. of them show affinities with neurotic classification. Detailed consideration of these would deflect from the main purpose of this paper and would more profitably be the topic of a separate paper when more empirical evidence has been accumulated, but some general remarks will be in order. In the first place, it may be noted that among the factors with significant neurotic correlation are U.I. 24 corresponding to the second order questionnaire anxiety factor, U.I. 32 corresponding to the second-order questionnaire introversion factor, U.I. 21 corresponding to the first-order questionnaire surgency factor. In addition to these we should mention Corticalertia (U.I. 22), Responsive Will (U.I. 29) and Competent Assertiveness (U.I. 16). These three are the most reliably associated with neuroticism and all seem to involve functions which one would expect to be worsened by neurosis. Cattell claims a multiple predictive correlation of the relevant objective tests with neurosis of .80. This is certainly impressive but not improbable if one remembers that the prediction is not in the sense of saying who will become neurotic sometime in the future, but who will be found to be already neurotic if examined by a psychiatrist.

The position so far is summed up in Table 1 where the inter-relation of the factors is shown graphically. It becomes obvious that there are only two important differences. In the first place Cattell has many more factors. In this respect he supplements Eysenck by his more precise and comprehensive analysis. In the second place the Cattellian objective factor U.I. 23 corresponds to Eysenck's objective neuroticism measures and five additional factors are needed for maximum prediction of neuroticism.

TABLE 1
EYSENCK

CATTELL

Questionnaire Factors :	Centroid { Neuroticism (emotional reactivity ?) Extraversion*	Second- { Anxiety (emotional reactivity ?) Exvia-invia
Objective Tests :	Neuroticism (found by direct correlation and criterion analysis).	U.I. 23 neurotic regressive debility. plus U.I. 24 as Anxiety above. U.I. 32 as Invia above. U.I. 29 Responsive will. U.I. 16 Competent assertiveness. U.I. 22 Corticalertia.

* Orthogonal and, therefore, necessarily not correlated with neuroticism.

Extraversion-Introversion v. Exvia-invia. Jung's classification of personality into two primary types is perhaps his most substantial contribution to modern psychology. Everybody seems to agree that this is an important concept but not all are agreed on the content of the concept. Jung (1941) himself stressed the inhibitory aspect. The control of behaviour and sociability (or its lack) was regarded as a direct outcome of this. McDougall too, seems to have had a similar attitude. His description of the action of depressant drugs in relaxing introversion depends essentially on the notion of reducing the inhibitory control of higher centres. This is the tradition which Eysenck has followed. It ties up well with Pavlovian theory and makes introversion-extraversion a product of basic psychological structure.

Cattell (1963) bases his corresponding concept of invia-exvia on the evidence from his 16 PF test which yields a major second order factor with many of the attributes usually assigned to introversion extraversion. High loadings on factor A (warm sociability), H (lack of shyness), E (dominance), Q2 (group dependency) and F (surgency) make it difficult to regard this factor as anything other than the traditional extraversion. Carrigan (1960) has made an excellent survey of factorial investigations in this area and shows that several researches using 16 PF factor scores and the Guildford Questionnaires agree in finding two factors which might contend for the title of extraversion-introversion. The first of these corresponds to the Cattell exvia-invia just described. The second has loadings on T (thoughtfulness) R (restraint), G (super-ego strength) and -F (surgency) and suggests a Hamlet like syndrome with inhibition at its core. This is probably a good match for one of Cattell's objective factors U.I. 17 (timid inhibition).

In the light of this discussion it might be suggested that we have two major introversion syndromes and that these represent two generalization effects. The Jungian's version relates to all the inhibitory tendencies ranging from timidity to ego and super-ego control. Timidity probably represents the chief genetic component. This same timidity is quite conceivably the source of an early and intense super-ego development. The child strives to please parents in order to maintain a high level of parental security. So he comes to introject

the parental code (Cattell G factor). At the same time he learns to avoid uncontrolled outbursts which may endanger parental approval and so develops strong ego-control (Cattell Q3 factor). Rather than *introversion* it might be better to think of this trait as *inhibitedness*.

Cattell's version of introversion represents generalization in another area. To a large extent the human being is engaged in relating to a social rather than a material environment. A variety of factors operate to encourage or discourage his seeking of social contacts. We have already noted the Cattell first-order factors involved. It is interesting to study the actual intercorrelations of these. With two exceptions, they form a consistent cluster. The first of these is that of dominance with cyclothymia: the warm friendliness of the latter conflicts with the cold ruthlessness of the former. The second exception is that of dominance with group dependency. Obviously the dominant person wishes to make the group dependent upon him rather than be himself dependent. Yet, despite the fact that dominance has conflicting effects with regard to these two traits, it obviously has the general effect of promoting social contacts since it counteracts any fear of social relationships. The dominant person is not shy but be it noted that the person who is not shy is not necessarily dominant. The traits are distinct though over-lapping in their effect.

We may ask whether inhibitedness and invia have anything in common. So far we have been concerned with their differences but it seems theoretically likely that the fear drive would provide a common element. Unfortunately, where these two syndromes have appeared in the same data, orthogonal rotation has been used and any possible correlation thereby obscured.

In collaboration with his wife, Eysenck (1963) has attempted to throw further light on this problem. They administered a questionnaire of seventy items, chosen to represent the social and inhibitory aspects of extraversion together with neuroticism, to 300 subjects. The principal component factors were rotated to orthogonal simple structure but still retaining a common extraversion factor spanning the bipolar sociability/impulsiveness items. It is argued that this is preferable (in terms of economy) to the logical alternative of two group factors of sociability and impulsiveness linked by a second-order factor.

This problem of deciding between a general factor plus a bipolar as against two correlated group factors is an old one. We would argue that it can be decided on grounds of parsimony but only when a very wide view is taken. Scientific parsimony applies only to the total system. If one has to choose between five well established factors or two new ones the former will be preferable since the latter involves a total of *seven* factors *versus* five for the former in the gross. The second-order factor may be a desirable concept if it results in a more meaningful whole, i.e., better integrating the specific case.

It is at this stage that we would invoke the principle which the Eysencks have so staunchly espoused: that factor analysis must be supplemented by more general hypothetico-deductive methods. We would suggest that the genetic inhibitory tendency postulated by Eysenck be recognised as operative in both the factor of inhibitiveness and the factor of invia but more particularly in the former. We would plead for recognition of the fact that the coherence of both these factors depends not merely upon a central causal factor (inhibition) but also upon the contribution of various functions to a common goal. For example, timidity, low emotional reactivity and high ego-control may all contribute to highly controlled behaviour without necessarily having a common causal element. It is this interplay of causes and effects which has to be disentangled by our scientific methods and neither factor-analysis nor any other method must be too rigidly applied. The ultimate aim is continually to attempt

simplification of our conceptual system and test widely the validity of such simplification.

In the present case we would argue that the total evidence favours an interpretation in terms of two positive factors of impulsiveness (or in the negative, inhibitedness) and sociability, correlated via the common genetic inhibition. The degree of correlation is possibly much less than indicated by the empirical test of correlating the factor measures. It is, in fact, possible to rotate the introversion and sociability/impulsiveness factors of the Eysencks into two separate factors of sociability and impulsiveness having all loadings positive and almost uncorrelated. We would consider a correlation of between .1 and .2 as being factorially indicated if one discounts the influence of the neuroticism factor.

From this study we may conclude that there is probably some correlation between sociability and impulsiveness but we are still very much in the dark as to the relation between the two introversion factors as distinguished by Carrigan and scored through the actual items from which they were originally derived. Nor do we know just where Eysenck's own neuroticism falls with regard to these two factors. Since he based his concept of introversion on that of Jung, it could be assumed that his own test would measure the Jungian type factor. In the light of the analysis just reported, we might conclude that it more possibly measures a high-order factor involved in both but this depends on the initial assumption that the data did adequately represent the two introversion factors concerned and it is possible that Carrigan's characterization of the two factors is an oversimplification.

A study by Bendig (1962) provides interesting material with regard to both these questions. The Guilford Zimmerman Temperament Survey which provides the major data for the Carrigan suggestion of two extraversion factors was analysed in conjunction with the M.P.I. scales. Two distinct extraversion factors emerge: one in the Guilford tradition and conforming to the Jungian concept the other characterized by sociability and ascendance. It is this latter factor on which the M.P.I. extraversion loads (.81 and .69 in the two samples). The former factor give insignificant loadings (— .12 and — .17). The two factors themselves turn out to be only slightly correlated (— .19).

Unfortunately, this evidence is based on a very small sample ($N = 54 + 58$) but it is so unambiguous and so similar in the two samples that it seems difficult to avoid the conclusion that Eysenck's extraversion, as measured by his test, is in fact, very similar to that of Cattell. A paper by Becker (1959) provides further confirmation. It includes the 16 PF second order extraversion factor and the M.P.I. extraversion score. Both of these get high loadings on a common factor (.82 and .68).

In the light of this discussion, it would seem that we must regard Cattell and Eysenck as being in substantial agreement in their measures of extraversion, despite the different theoretical backgrounds—social inhibition and general inhibition respectively—from which they may have approached.

IV.—PSYCHOTICISM.

We have so far made no reference to Eysenck's third factor: psychoticism. This did not appear in his original research, but emerged from special studies designed to show whether psychotic disorder is on the same continuum as neurotic disorder (1952), as the neuroticism factor involved all that the psychotics had in common. The evidence indicated a common psychoticism factor which was distinct from that of neuroticism.

Cattell used similar tests in his O-A Battery and found a first-order factor on which they had high loadings: U.I. 25 (Realism *versus* psychoticism).

This confirms that psychoticism as measured by objective tests, is similar on both sides of the Atlantic, but still leaves open the possibility that, as in the case of neurosis, some other factors may be found to be important. The chief aspirant here may be Cattell's L (paranoid), which is the only 16 PF factor on which psychotics differ considerably from normals but since the score is very low, it may be suggested that this is just another indication of their lack of reality orientation and a manifestation of the same trends which come out in the psychoticism tests themselves.

The aim of this paper has been to highlight the basic agreements and disagreements in the findings of these two factorists. It seems clear that the major differences arise from methodology and that even these are much less than might be expected. The most striking fact is probably that there are no findings which cannot be reconciled. Had Eysenck continued his factorial studies, intensively, the two systems would probably have come into accord. In the interest of reducing confusion and by way of summary of this paper, the following guide is offered:

- (1) For Cattell's *Anxiety* and Eysenck's *Neuroticism* read *Emotional reactivity*.
- (2) Regard Cattell's *exvia-invia* as substantially similar to Eysenck's *extra-version-introversion*, and neither as basically Jungian.
- (3) Match Cattell's neurotic regressive debility with Eysenck's objective tests of neuroticism.
- (4) Note Cattell's additional predictors of neuroticism.
- (5) Bear in mind that Eysenck's centroid factors correspond to Cattell's second-order factors, but do not permit the finer statistical implications of the latter.
- (6) Take account of the finer distinctions provided by Cattell's first-order factors which throw fuller light on the complex causal relations involved.
- (7) Look at Eysenck's work always in relation to his general hypothetico-deductive approach and the related conceptualization.

V.—REFERENCES.

- ¹ BECKER, W. G. (1959). *Cortical Inhibition and Extraversion-introversion*. USPHS Proj. No. M1713, Univer, Illinois.
- ² BENDIG, A. W. (1962). Factor analysis of the Guilford Zimmerman Temperament Survey and the Maudsley Personality Inventory. *J. Gen. Psychol.*, 67, 21-26.
- ³ CARRIGAN, P. M. (1960). Extraversion-introversion. *Psychol. Bul.*, 57, 329-360.
- ⁴ CATTELL, R. B. (1946). *The Description and Measurement of Personality*. Yonkers, N.Y. World Book.
- ⁵ CATTELL, R. B., SHEIER, I. H., and LORR (1962). Recent advances in the measurement of anxiety, neuroticism and the psychotic syndromes. *Ann. N.Y. Acad. Sc.* 93, 813-856.
- ⁶ GORSUCH, R., and CATTELL, R. B. (1963). *Confirmation and Precisioning of the Second-order Personality Structure of the 16 P.F.* (In press.)
- ⁷ EYSENCK, H. J. (1944). Types of personality: a factorial study of seven-hundred neurotics. *J. Ment. Sci.*, 90, 851-861.
- ⁸ EYSENCK, H. J. (1950). Criterion analysis: an application of the hypothetico-deductive method to factor analysis. *Psychol. Rev.*, 57, 38-53.
- ⁹ EYSENCK, H. J. (1957). *The Structure of Human Personality*. London: Methuen and Co.
- ¹⁰ EYSENCK, S. B. G., and EYSENCK, H. J. (1962). A factorial study of an interview questionnaire. *J. Clin. Psychol.*, 18, 286-290.
- ¹¹ EYSENCK, S. B. G., and EYSENCK, H. J. (1962). The validity of questionnaires and assessments of extraversion and neuroticism and their factorial stability.
- ¹² EYSENCK, S. B. G., and EYSENCK, H. J. (1963). On the dual nature of extraversion. *Brit. J. Soc. Clin. Psychol.*, 2, 44-55.
- ¹³ JUNG, C. J. (1941). *Modern Man in Search of a Soul*. London: Kegan, Paul, Trench, Trubner.
- ¹⁴ THOMSON, G. H. (1946). *The Factorial Analysis of Human Ability*. London: University of London Press

(Manuscript received February, 1964.)

CRITICAL NOTICE

WISEMAN, STEPHEN (1964). *Education and Environment*. Manchester University Press, pp. xiv+202, 37s. 6d.

By CYRIL BURT

The revolutionary changes in both social and educational conditions that have occurred since the war have been accompanied by much heated controversy, most of it, as Professor Wiseman points out, based on political prejudices rather than on ascertained facts. The decision of the Research Committee in the Manchester School of Education to undertake a large and systematic inquiry into "the relations between educational attainment and social and environmental factors" was thus both natural and well-timed. The present volume reports the chief results.

In planning the investigation it was resolved to concentrate primarily on the new age-group of pupils between 14·0 and 15·0, produced by the raising of the school leaving age. The area covered consisted of the boroughs of Manchester, Salford, and Stockport, together with the adjacent boroughs and urban districts in Lancashire; and the schools selected comprised all those maintained by the local authorities, except the special schools. Private or independent schools, and those receiving direct grants, were excluded. The number of children tested turned out to be nearly 14,000, of whom nearly 40 per cent. were in 'all-age schools,' 35 per cent. in modern, 15 per cent. in grammar, and the remainder in either technical or selective central schools. The subjects available for testing in these different schools were inevitably limited. Those chosen were reading for comprehension, mechanical arithmetic, and verbal intelligence. Tests suitable for age-group 14 had to be specially constructed and standardised.

The main research was, we are told, expressly designed so that the results would be comparable with the study of educational backwardness carried out in London between the two wars. A comparison of the figures shows that certain relations still hold, while others have been greatly modified by post-war conditions. Maps are printed to show how the distribution of backwardness in reading, arithmetic, and general ability corresponds with that of certain environmental characteristics, and correlations are calculated between the figures for the twenty-six wards. Among the various local characteristics selected for comparison in the earlier research the frequency of mental deficiency still shows the highest correlation with the incidence of backwardness in educational tests; slightly lower correlations are obtained with birth-rate, illegitimacy, and neglect; the correlations with the death-rate, with infant mortality, and with density of population, however, are now much smaller. These limited changes are, I think, precisely what we should anticipate as a result of post-war social and economic changes. A factor analysis reveals three main groups of determining conditions: (i) a 'genetic component,' (ii) 'economic background,' and (iii) 'lack of parental or maternal care.' "Economic background," we are told, "has little to contribute towards either backwardness or brightness." In particular, "environmental factors predict backwardness in reading with less precision than they predict backwardness in arithmetic; the concomitants of reading (parental interest and the cultural level of the home, for example), seem to be 'psychological' rather than 'economic' or 'physical'." 'Brightness' appears to be a more complex characteristic than 'backwardness': the two "are not just the obverse and reverse of a single coin." However, as Professor Wiseman remarks, the intentional limitations of the research render the findings in regard to brightness suggestive rather than conclusive; and he emphasizes the urgent need for "an inquiry of wider scope, embodying a still greater variety of environmental variables."

The account of this preliminary survey is followed by a chapter from Dr. Warburton, describing a further study of the relations between attainments and the

sharacteristics of the school—type of building, size of school, size of class, progressiveness, neighbourhood, etc. Detailed analyses were carried out—by correlations, by factors, and by variance. In the main the results confirm those of earlier investigators; but there are a few striking differences. Size of school is apparently not so important, as Kemp and others have inferred (see this *Journal*, XXV, p. 67); on the other hand, "progressiveness and smaller classes appear closely associated with higher attainment." But "the most surprising finding is that the correlations (with social conditions) are higher for intelligence than for reading or arithmetic, in precise contradiction to theoretical expectation." Dr. Warburton discusses this result at some length, and finally—though somewhat reluctantly—concludes that "on every count these findings give more weight to the environmentalist than to the genetic viewpoint."

However, the result is neither quite so new nor so unexpected as he seems to suppose. It has often been noted when correlations are based on groups (e.g., schools or districts) instead of on individuals, and appears to depend mainly on the policy of the teachers in the area. In the poorer schools within the area here studied, "the amount of time and effort given to basic arithmetic and reading," we are told, "is much greater than in the more fortunate suburban schools." Yet this, I suggest, is not the whole story. The reader can easily determine the 'theoretical expectation' if he takes the familiar formula for the 'total covariance,' $r_{sg} \sigma_s \sigma_g$ in the case of social status (s) and general intelligence (g)—with a similar expression for attainments, and analyses each in the usual way into the sum of the covariance 'between schools' (i.e., between the means) and the covariances 'within' schools. It then becomes clear that the reversals which have puzzled Dr. Warburton and other investigators are produced, not by the mere reduction of deviations in the attainments, but by changes in the correlations within schools. And this, in turn, suggests that, if instead of striving to improve the level in the weaker schools, the policy is rather to improve the attainments of the weaker *individuals*, then, so far from the school correlation for attainments being diminished, it will be increased. And that, in fact, is what we commonly find.

Professor Wiseman himself lays special stress on the need to preserve the 'genetic viewpoint.' "We are," he says, "living in a period of egalitarianism which carries with it a hostility towards the concept of differences in innate ability: . . . many enthusiastic supporters of comprehensive schools reject completely the objective tests of aptitudes and attainment, which, in (his) view, are essential tools in efficient organization. No doubt, the pendulum will slowly swing back. Meanwhile, how many of our pupils will suffer?" But throughout the volume Professor Wiseman is careful to give full weight both to heredity and to environment, emphasizing this factor or that with scrupulous impartiality according to the balance of evidence in every case.

I have had space only to touch upon a few of the many important inferences that he has drawn from these admirable surveys. And the whole book should be closely studied, not only by those contemplating similar investigations, but by every teacher, by every educational official, and I would add, every student who is intending to become a teacher.

BOOK REVIEWS

ABERCROMBIE, M. L. J. (1964). *Perceptual and Visuomotor Disorders in Cerebral Palsy*. Heinemann, pp. 136, 25s.

This book is the outcome of a five-year study, sponsored by the National Spastics Society, with the aim of bringing together the recent work on this topic, and clarifying the present position. The author not only brings together much of the work, but also highlights many of the difficulties and causes of apparently conflicting results. Gaps in our present knowledge are shown up, revealing the most urgent areas for further research.

The book begins with a discussion of the nature of the disorder. A dichotomy of 'perceptual' and 'visuomotor' is adopted, as a preliminary attempt to sort out the confused thinking and vague terminology which exists at present. The need for more work on the normal developmental stages of perceptual and visuomotor skills, and for well standardised tests is pointed out.

The specific nature of the disabilities is discussed, and also the suggestion that these may be a developmental lag as opposed to a distortion. This first chapter gives a brief, clear introduction to the present position, and is fully illustrated with diagrams and line drawings.

There follows a chapter on the incidence of the disorder and its relationship to other handicaps. The incidence in different types of cerebral palsy is discussed, and the possible effects of motor and visual handicaps. This aspect of the subject is bedevilled by differences of categorisation with regard to type and extent of handicap, and by criteria for specific disabilities. It is only possible to draw the very broadest conclusions, and the author is wisely cautious.

In a useful chapter on "the strategy and tactics of research," methods are grouped as (1) clinical and (2) statistical and actuarial. These two approaches are fully discussed, with examples. As a critical review of research methods, this chapter should provide much food for thought for anyone attempting research in the clinical field.

Throughout these first chapters the need is constantly brought out for more normative studies and for more research into educational implications.

Chapter five is an attempt to cover a wide range of tests of perceptual and visuomotor abilities in a general way, with specific examples. The tests are grouped according to the particular area of functioning on which they depend, within the broad dichotomy of perceptual and visuomotor functioning. This chapter should be a useful reference list for anyone planning research on any aspect of this subject, but it is a little sketchy if it is intended for a wider reading public. Fewer, and more fully discussed examples might have been more generally useful.

The second half of the book consists of abstracts of relevant work, arranged alphabetically under the authors' names. These abstracts are clearly set out and headed so that they may be used for quick reference.

We should indeed be grateful to Dr. Abercrombie for bringing together, so clearly, this mass of information and references. The book should have value as a well balanced introductory coverage of the subject for psychologists, doctors, and the more interested specialist teachers. It should also be a guide and stimulus to further reading and research for psychologists particularly concerned with this aspect of intellectual functioning, from the genetic or pathological point of view.

JOAN REYNELL.

BATTIN, R. RAY, and HAUG, C. OLAF (1964). *Speech and Language Delay*. Illinois: Charles C. Thomas, pp. xi+71, \$4.50.

The subtitle of this book describes it as a "home training program." The authors are both highly qualified and experienced in the field of hearing and speech pathology and have worked with the handicapped. Here, they are addressing themselves, not to other specialists, but are writing a practical manual for parents.

The book opens with a review of normal early speech development. This is on expected lines, though some of the more general statements (for instance, on sex differences) require qualification. There follows a chapter on the contribution of home discipline and after this comes the 'educational program.' It is in four rather arbitrarily divided sections devoted to stimulation, motivation, ear-training and auditory memory training, respectively. Closing chapters deal with helping the hard-of-hearing child (this includes some useful advice on the care and adjustment of aids) and finally the parent is considered in the role of teacher.

The merits of the book lie in its simplicity and clarity. Important practical points are brought out with unmistakeable emphasis: "Talk, talk, talk, to your child." Definite hints and suggestions, with photographs, help to get the methods across. There is obviously much sense and a back-ground of practical experience in what the authors advise. If lack of research evidence for their views bothers the academic, this would probably not seem a fault to the parents.

But having a parent readership so obviously in mind, it is rather surprising that the writers do not emphasise the varying needs of each individual child more. Even the differences in approach required by the mentally backward as compared with the deaf, are not adequately faced or explained. This could result in frustration of the efforts of the conscientious parent-reader, who would otherwise find the book a valuable guide.

O. C. SAMPSON.

BLAIR, G. M., and JONES, S. R. (1964). *Psychology of Adolescence for Teachers* (pp. 118, 11s. 6d.).

CHARLES, D. C. (1964). *Psychology of the Child in the Classroom* (pp. 86, 9s. 6d.)

STRANG, R. (1964). *Guidance in the Classroom* (pp. 118, 11s. 6d.).

All in the Psychological Foundations of Education series, editor V. H. Noll.
New York: Macmillan; London: Collier-Macmillan.

The present fashion in America seems to be to replace the small numbers of massive textbooks prescribed for students' use by large numbers of small paper-backed volumes, each covering a specific topic. There is a lot to be said for the practice; many textbooks cover topics not needed by all students; students can each buy several different paper-backs which can then be shared, they are easily transportable and, as the system spreads, there are more choices open to students who may study similar subjects in different series.

The three titles in this new series are written at an appropriate level for students training for teaching. In each, the field is fully covered and supplemented by a list of well-chosen, recent references to primary and secondary sources. There are no indices, but perhaps with such short books this is not a serious drawback.

British students will undoubtedly find these books interesting and informative; the only difficulty lies in their essentially American orientation which applies not only to all the references, but also to the background material used in illustration. This is seen particularly in the book on Adolescence, where the norms of behaviour and often the form it takes are so different from our own, that our students might be seriously misled if they relied on this text alone for a picture of teen-age behaviour. At the same time, the account here could make a good starting-point for seminars comparing the two cultures.

In all three books the emphasis is upon practical classroom management, but important theoretical aspects are given due place. With the reservation already stated, they can be recommended for British students in colleges of education, or as preliminary reading for more advanced students of educational psychology.

L. B. BIRCH.

BRAUNER, CHARLES J. (1964). *American Educational Theory*. London: Prentice-Hall International, pp. ix+341, 48s.

This scholarly and well-written book provides a searching commentary on six major traditions of educational thought in America. In chronological order these are: (1) the Bell-Lancasterian monitorial system with its emphasis on strict obedience, verbal drill and mechanical learning; (2) Pestalozzi's object-teaching conceived within a framework of affection and individual attention; (3) the Herbartian view of education as a rational science based on ethics and psychology; (4) the child-study movement which strove for 'felt understanding' through the direct observation of behaviour; (5) experimentalism, as exemplified in the problem-solving, activity and project methods of Dewey, Kilpatrick and Monroe; (6) the current academic emphasis which, according to the recent Rockefeller and Conant studies, court the virtues of balance, scope and high levels of achievement. In respect of the latter, Dr. Brauner says, "Conant gives fully 80 per cent. of his attention to the talented and the gifted college-bound student." What is more, "For all the talk about individual courses of study, scope becomes a matter of fitting the student to the academic curriculum rather than fitting a curriculum to his capacities. The 'interest' and 'ability' employed in guidance are determined by tests of scholastic aptitude, the recorded achievement as measured by grades in courses, and by teachers' estimates'" (p. 276). Great balls of fire! The Americans have at last exceeded the English in creating and perpetuating categories of failure within their educational system!

Dr. Brauner analyses each of the above-mentioned traditions as a general method of teaching and as a subject of academic study and concludes that "both general methods and general disciplines continue in business long after their practical uselessness and their intellectual unsoundness have been pointed out. General methods and general disciplines hold their shape like sucked-out egg shells: kept rigid by the tension of fictional hypotheses supporting poetic analogies" (p. 299). But criticism does not stop at the journalistic bias in American educational thought. The author also deprecates the prevailing demand that ideas must be made immediately practical—a tendency which "so far outweighs new information that it makes some sense to say that education as a subject of study does not exist" (p. 302).

Among the remedies suggested by Dr. Brauner for improving this state of affairs is a plea that increased emphasis be given to conceptual research which he sees as "the missing link between a logical-speculative tradition prematurely abandoned and a body of immediately useful facts without theoretical portent." (p. 4). In this, as in his other observations on such matters as the professional training of teachers and the value and limitations of measurement research, Dr. Brauner's arguments are clear, consistent and well-reasoned. Altogether then, this book provides not only a stimulating discussion of the nature of education as an intellectual discipline but of the means whereby its academic and professional aspects may become more closely integrated. It is a scholarly performance characterised throughout by wit and good sense and is to be strongly recommended as a most useful breviary for all lecturers in colleges and departments of education. There is, incidentally, an interesting appendix which includes a hitherto unpublished memorandum by Dewey outlining the essentials of a modern school of education. Topical if you like! E. G. S. EVANS.

COOMBS, C. H. (1964). *A Theory of Data*. John Wiley and Sons, pp. xviii+585, £5 13s.

Regular readers of Coombs will recognise many of the elements of the theory of data presented in this volume; but they will, I think, be excited by the new framework. In the first of six parts an overview of the theory is given. Data are seen as relations between points in a space. The types of relations which may hold between two distinct sets of points—for example, individuals and stimuli—are systematised, symbolised and axiomatised to a high degree of generality. This leads to a fourfold classification of data as preferential choice, single stimulus, stimulus comparison and similarities data, which are discussed extensively in Parts 2 to 5.

A final section reviews the profusion of relationships between the elements of the new theory and, in a postscript, reminds us of the old theory.

The familiar Unfolding Theory, in which both individuals and stimuli (or any two sets of points) are found to be in a joint space, appears in many forms throughout the book with ample exposition and illustration on each occasion. This and the repeated explanation of the types of data palls, but since readers will probably dip into the book the repetition is a useful feature. A thorough reading of Part 1 is, however, essential to an understanding of the remainder.

Throughout, familiar kinds of data are given new perspective. The whole is embedded in a clear statement of the role and responsibility of the scientist: he selects the behaviour to be observed; he classifies the observations as belonging to a particular data mould; and he chooses a model for "the detection of relations, order and structure which follow as a logical consequence of the data and the model used for analysis" (p. 4). A repeated statement is that "we buy knowledge with the assumptions we make—the more assumptions made the more knowledge obtained" (p. 284). In spite of this, Coombs is not disposed towards making too many strong assumptions. He patiently exhausts all possibilities generated by a few, relatively unrestrictive, assumptions before he proceeds to harsher but more robust methods like probability arguments and factor analysis. Clearly, Coombs is essentially deterministic in his approach: discrepancies in the data are first taken to indicate the need for a modification or extension of the model rather than the need to assume error components and some overworked statistical model. This is, indeed, refreshing.

This volume may be regarded as complementary to Torgerson's *Theory and Methods of Scaling*. Although there must be a high degree of overlap in their references and their content, Torgerson's contribution is the grand review whereas Coombs' is the grand revision. For many years to come, both are essential reading.

PHILIP LEVY.

EYSENCK, H. J. (1964). *Crime and Personality*. London: Routledge and Kegan Paul, pp. xv+204, 25s.

From its title this book might be expected to be peripheral to the interests of readers of this *Journal*. This, however, is not true for, although it is mainly concerned with delinquents, the basic problem is that of bringing about changes in human behaviour in socially desirable directions. Educators have always regarded this as one of their functions though, as Professor Eysenck points out, in manipulating people, it is doubtful whether our educational systems are any more successful than those which prevailed in earlier times.

The book starts with the premise that present-day views about crime and criminals are largely notional and discount the considerable amount of verified psychological knowledge and of proven techniques which make it possible for at least a start to be made towards a scientific study of criminality for the purpose of discovering causes and then means of treatment. What follows is a natural development of Professor Eysenck's own theory of personality which was expounded at length in his *Dimensions of Personality* and *The Scientific Study of Personality* and which is succinctly summarised here in its relevant aspects. Evidence is cited to support the view that heredity plays an important part in predisposing certain people to become delinquent. This comes through their being much less susceptible than most people to the social pressures which give rise to the development of conditioned fear responses to behaviour which is disapproved of by parents, teachers and others in authority. These poor conditioners fail to develop a strong 'conscience.' Two approaches to the treatment of criminals are suggested as promising lines of research. The first resembles behaviour therapy in that it involves the application of modern learning theory in the light of critical diagnosis of the criminal's position along the significant dimensions of personality. The second suggestion is that

appropriate drugs should be used to modify the individual's deviant personality in the direction of greater introversion and easier conditioning so that a conscience may develop.

Professor Eysenck writes very well, his arguments are convincing, his experimental evidence well and fairly chosen, and he frequently exposes the links in his logical chain which are weak or missing. What he does is present a vulnerable hypothesis and a challenge to research workers to test it. If the hypothesis proves substantially true, it then demands a revolution in many of our legal and therapeutic procedures with delinquents. Inevitably the ideas expressed here will provoke opposition. This will be good so long as the opposition is not merely verbal argument but well-founded experiment.

This, in spite of the ease with which it can be read, is not perhaps a book for 18-year-old education students, but for lecturers and mature members of advanced courses it should be essential reading, as it should be for magistrates and others interested in penal procedure.

L. B. BIRCH.

GALLAGHER, J. R., and HARRIS, H. I. (1964). *Emotional Problems of Adolescents*. London: Oxford University Press, pp. 210, 35s

This new edition of *Emotional Problems of Adolescents*, revised and with some useful new material on adult-adolescent discussion, and on the quandary of adolescents in a materialistic society, comes at an appropriate time. However tardy at first, those in charge of secondary education in Britain are now beginning to act on the well-established fact that an adolescent community throws up enough personal problems to keep any available counsellors constantly busy. With this goes the correlate that the amount of guidance needed cannot be provided by the skilled personnel in the Guidance Clinics and School Psychological Services alone, but must also be available through friendly adults endowed with the right personality, and equipped with a dependable framework of knowledge and approach.

The authors, one a pediatrician; the other a psychiatrist, make this situation their starting point. They regard general practitioners, teachers, club leaders and clergymen as today carrying the chief responsibility for helping adolescents, as a supplement to whatever parents can do, to surmount the various difficulties likely to be encountered in the individual struggle towards maturity. The authors set out to share, with all those intimately concerned with adolescents, both what they have found to be the fundamental problems of adolescents and also their experience in dealing with such problems. The result is the communication of a wealth of insight through analysis, precepts and case histories.

Cultural differences between the American and British environments inevitably emerge and are of considerable interest in themselves. It would seem that mother dependence among adolescent boys is a good deal more common on the other side of the Atlantic than in Britain, thus warranting a separate chapter on "Home-sickness." The ambivalent attitude to sex in America today—*pseudo-uninhibited*—gives a rather old-fashioned slant to some of the comments on adolescent sexuality. Status-seeking comes through as an even more intense source of stress among the young in America than in Britain.

Certain theoretical inconsistencies obtrude a little. At times the book would seem to have a Freudian orientation; at others it ranges well beyond the limitations of the psycho-analytical viewpoint. This produces occasional curiosities, as when, for example, masturbation among adolescent boys is described as a regression. But the wisdom of the book overrides such anomalies. The chapter on adult-adolescent *tete-a-tete* discussion is probably the best succinct treatment in the literature.

JAMES HEMMING.

GARDNER, D. BRUCE (1964). *Development in Early Childhood: The Pre-School Years*. New York: Harper and Row; London: Evanston, pp. x+358, 43s.

This book, by Professor Gardner, of Iowa State University, provides a companion volume for *Behaviour and Development from 5 to 12* (from the same publisher, 1962) by Glenn Hawkes and Damaris Pease, also of Iowa University.

The book is in four parts, How We Study Children, Foundations of Development, Aspects of Development in the Pre-School Years, and the Society of the Pre-School Child (mainly family and nursery school).

The author has encompassed a very great deal of matter within a relatively short book, in a masterly fashion which could be attempted only by a skilled teacher who is experienced also in child guidance. In the first two chapters, for instance, he summarises the history, methods, viewpoints, practical relevance and scientific methods of child development (p. 79). He identifies four major streams of thought, "the efforts of such men as Watson, Gesell, Freud and Lewin," as having laid the foundations for a new concept of the nature of children. This contemporary viewpoint he develops throughout the book as a child-centred one, in which the function of adults is "to provide a world of objects and a world of people that actively assist the child in his efforts toward self-realization" (p. 349).

An attractive feature is the provision of a few questions at the head of each chapter, cleverly arranged to stimulate intellectual curiosity and also to turn attention at the end to practical considerations. For example, the four questions at the beginning of chapter 7 (Communicating with others) begin with "What differences are there between learning to speak and learning to communicate?" and end with "What specific tasks must the child accomplish in order to communicate well with others?" The author is specially good at putting issues into perspective, and he does this partly by selecting main points, for summary at the conclusion of each chapter.

This book would be useful for Colleges of Education, and for first-year courses in Child Psychology at Universities, with the possibility of further study through the references given at the end of each chapter. There are twelve diagrams or tables, and a number of photographs. It is intended also for students of Home Economics in America, where the subject is interpreted more broadly than in most of our courses of Domestic Science, to include the psychology of child rearing. L. R. BOWYER.

GREEN, EDWARD J. (1963). *The Learning Process and Programmed Instruction*. Holt, Rinehart and Winston, Inc., pp. xii+228.

AUSTWICK, K. (Edit.) (1964). *Teaching Machines and Programming*. Pergamon Press, pp. vii+205, 42s.

That teaching machines can be useful tools there is no denying, and in some subjects their use by groups of pupils, for restricted periods of time, may be profitable. These subjects include mathematics, physical sciences and foreign languages. The machines have also established their usefulness for training in industrial skills, and for tuition in the Armed Forces. Perhaps the greatest benefits that this movement has begun to bring, however, is a reconsideration of teaching methods and of the writing of text books, and a re-emphasis of the importance of individual rather than group learning. Perhaps its greatest dangers are of a certain de-humanisation of learning, and the lack of a sufficient realisation among some of its exponents that the learning of facts, of skills and even of principles and methods, is not by any means the whole of education. Nor has this reviewer seen anywhere a perceptive appraisal of the possible effect upon a child of being required to conduct a large part of his learning via teaching machines. The human mind needs variety not only of subject matter but also of teaching method, and the soundness of a method which gave the child long periods at teaching machines, even with rest pauses, would be questionable.

The book by Green is thoughtful and well reasoned. In the first part on the learning process, however, the author is always writing in abstract terms, without providing illustrations of his points. For this reason, although those who persevere will be sufficiently rewarded, there will be many, even among psychologists who will soon put the book on one side. This is unfortunate, not only because the fault could have been easily remedied, but because many will thereby miss the excellent chapter on techniques of programming, and the useful one on evaluation. In this section also the author has some very good comments on such practical problems as cost.

The second book is very much easier to read. Refreshingly lucid, in contrast to the American book, is Kay's substantial 'General Introduction,' which explains and critically evaluates the principal systems. The editor contributes one of the chapters and there are six contributors in all. The topics include: 'Teaching machines in industrial and military training,' and 'The adaptive teaching system.' Inevitably there is some overlapping, and naturally those contributors writing at the end suffer from this more than those whose contributions are read first. The price is atrociously high—unless the extra goes to the writers!

R. R. DALE.

LAWSON, C. A., and BURMESTER, M. A. (1963). *Programmed Genetics, Vol. I, The Basic Concepts*. Boston: D. C. Heath and Co., pp. iv+329, no price given.*

Many students of the biological and social sciences lack specific biological training yet need more than passing acquaintance with genetics. It is more difficult to make up a deficiency in early training in genetics than in most subjects and students will undoubtedly find this programme of great value.

It is a scrambled text book; straightforward in lay-out and easy to work through. Unlike some of the scrambled texts which have been produced, it is also thoroughly adult in language and approach. The problems to be encountered are stated beforehand so that one's progress is signposted. The steps in the arguments are clearly developed in a most illuminating way, and the assumptions on which they are based are given. Each chapter ends with a short test and a summary which provides a useful revision exercise when a period of time elapses before commencing the next.

It is inevitable that this should happen as the programme is designed in six units, each of which should occupy one period of work. The average time taken by American students to work through the whole programme is given as fifteen hours, though I would think that this is a generous allowance. Those who are already familiar with some genetic concepts may find progress slow but those who struggle on in the dim light of semi-understanding will find the time spent rewarding. However, it would probably speed things up and be a considerable help if the digrams, to which one has constantly to refer, were placed at the end of the book on folded sheets of paper so that they could be extended and kept in front of the reader the whole time.

Part of the programme has been tested and revised on 144 students who did better on a test than a class taught group of comparable size and ability. More information concerning the evaluation of the programme is available in *Science Education*, 1960.

The authors are to be congratulated in producing a text of such high standard and in leaving us to look forward to Volume II.

P. K. POPPLETON.

* * * * *

This study programme set out to teach basic genetics without recourse to formal lectures. It assumes no prior knowledge other than the ability to read. Within its self-imposed limits it succeeds admirably.

*As teachers of psychology often need to teach basic genetics to their students it was felt desirable to have this book reviewed by a psychologist (P.K.P.) and a geneticist (R.A.W.). Editor.

The student is introduced to the principles of genetics by a consideration of human variation and a family pedigree of a particular trait. This is probably more suitable than the conventional methods with peas or fruit flies since the student can see such variation in his fellows. The second chapter deals with pedigree analysis at the genotypic level and sets out the principles of the gene theory. Crosses between pure lines of animals and plants are introduced in chapter three and the results of these crosses are correlated with the gene theory.

Probability, a concept which many students find difficult to grasp, is treated well in the fourth chapter. The transition from tossing pennies to working out genotypic ratios is hardly noticeable. I did find the authors method of expanding probability equations rather ponderous and on page 168 they introduce the term 'binomial' without explaining what it means.

The chequer-board method of calculating genotypic ratios is introduced in Chapter IV. The sixth chapter covers the inheritance of two pairs of contrasting characters in three stages: first the principle of independent assortment, second the chequer-board analysis of a dihybrid cross and finally deviations from anticipated ratios due to various types of gene interaction.

The test sections at the end of each section and chapter are extremely searching. It would be more convenient if the answers did not have to be written in the book, this particularly applies to the chapter reviews.

There is one typographical error of some importance, in page 182A the third possible solution should read P and p —the p was missing in the copy I had. This solution is unfortunately the correct one!

In all parts of the programme the treatment is extremely thorough. Its scope is rather less than that specified in G.C.E. "A" Level syllabuses. However, as one who has had to counter the hazy and often misleading ideas of genetics that students pick up at school, I would recommend it not only for schools but also as an adjunct to first-year university courses in biology.

R. A. WOODS.

LEVIN, A. KATHARINE (1964). *Cerebral Palsy—The Pioneer Years of Occupational Therapy in Scotland*. Edinburgh and London: E. and S. Livingstone, Ltd., pp. xi+120, 30s.

The amount of information given in this book is in inverse proportion to its brevity (56 pages of text supplemented by 98 photographs).

Written primarily as a record of the first fifteen years of Occupational Therapy at Westerleigh School, Edinburgh—the first O.T. service established in Scotland specifically for cerebral palsied children—the book describes the gradual expansion of the concept of treatment from 'gait and posture sessions' to the inculcation of increasing social self-reliance and what Dr. J. A. L. Naughton so aptly terms in his Foreword "... a minimum of blinkered egotism."

The methods used were basically those of the late Dr. Temple Fay but, with the expansion of the service, originally for resident pupils, to out-patient children and adults, application of these methods was extended beyond the confines of school and treatment room, via a 'cottage workshop' to a public laundry service operated by ex-patients and recognised by the Ministry of Labour.

This latest venture (begun in 1961) is of particular interest in providing not just a subsidized sheltered employment but, to those unable to obtain, or keep, employment in the open labour market, an opportunity to help start and run a business which, after an initial loss, soon began to show a small, but steadily increasing profit.

The account is written clearly and simply in a straightforward manner. The style is factual, uncluttered by emotionality or over-long case histories, reference to individual patients being used only where necessary to illustrate particular problems or ways of dealing with them. The photographs have been well selected to illustrate specific points made in the text; the index is arranged to provide quick reference to the photographs illustrating textual descriptions and there is a short bibliography. The printing and general format is excellent.

This work is recommended to all who work with cerebral palsied persons. Not only has Miss Levin demonstrated that the ultimate value of occupational therapy cannot be realised fully unless opportunities are given outside the treatment room for the wider application of skills acquired therein, but she has, incidentally, written a testimony to the initiative, enterprise and personal adaptability of all the Staff (herself included) who were concerned in the work described. M. I. DUNSDON.

MELTON, ARTHUR W. (Edit.) (1964). *Categories of Human Learning*. New York, London: Academic Press, pp. xvi+356, 68s.

Early in 1962, a Symposium on the Psychology of Human Learning was convened at the University of Michigan. Its theme: the "inter-relatedness of different categories of human learning, with emphasis on (a) the formulation of statements about the categories that reflect the sophisticated methodology of contemporary laboratory studies, and (b) the relatedness of the different categories of human learning—whether in laboratory practice, in empirical generalisations about organismic or procedural variables, or in theory." With a slight change of title, and opportunity for some second thoughts, this book presents the papers of that symposium.

As the statement of the symposium's theme promises, this is no popular explanatory text. None of the contributors, each a specialist in his own field, makes more than passing concessions to the reader who is not. The book is, nevertheless, well documented, the papers are lucid though highly condensed, and the psychologist reader with strong motivation will be rewarded for his efforts. Amid so many oversimplified accounts of human learning, one can welcome a book of this high-level kind.

For each of the seven topics discussed there is a main speaker and a commentator. The topics are: Classical and Operant Conditioning; Rote Verbal Learning; Probability Learning; Short-Term Memory and Incidental Learning; The Concept of the Concept; Perceptual-Motor Skill Learning; and Problem Solving. Professor Melton, as editor, provides a final overview.

It would be impossible to review the content of each of these tightly-packed specialist papers, and invidious to select. One may, nevertheless, comment on the extent to which the book as a whole fulfils the aim of dealing with "the definitional and taxonomic issues that plague the psychology of human learning . . .", or at least of focusing the attention of the seven principal participants upon them. Despite the editor's warning that "it was not expected that the principal participants and their discussants would provide psychologists with a new taxonomy of human learning . . .", one may, nevertheless, feel somewhat disappointed. What each of the contributors certainly has done is to demonstrate the complex and increasing proliferation of his own specialism. To take but one example: far from being offered a higher view embracing both classical and operant conditioning in one scheme, we now have four subcategories of the former and eight of the latter. Nor, indeed, is there unanimity about the crucial *differences* among the categories. Such proliferation is a natural consequence of the development of more refined experimental techniques; to be expected, and in many respects welcomed; but it does not of itself contribute directly to the systematic organisation of our knowledge of human learning.

A hint as to how the latter problem might be tackled is contained in a remark by Professor Estes. "The various familiar types of learning are the *result* (reviewer's italics) not of any rational scheme of classification, but simply of the clustering of research activity around points of communality in problems and procedures." In short, the categories of human learning which are to be related are themselves theoretical, and not sacrosanct data from which further abstractions must be drawn. In contrast, some of the papers appear to presuppose that (so to speak) there 'really are' two basic kinds of conditioning, three categories of this, four of that, and so on; and that the taxonomic problem consists in discovering features common to these categories. Here and there throughout the papers, and particularly in the

editor's final comments, it is suggested that the problem is essentially a conceptual one; but rarely are the implications of that view taken up seriously. As Professor Melton almost says, we can no longer excuse ourselves on the ground that we need further experimentation and more facts. What the book makes abundantly clear is the need of new concepts, which will not merely find common ground in the now traditional categories but redraw the map. But it also makes abundantly clear that the thinkers who tackle that need are faced with a Herculean task; its content issues a clear warning against overoptimism.

JAMES M. THYNE.

MUSSEN, PAUL H. (1964). *The Psychological Development of the Child*. Prentice-Hall, Inc., pp. 109, 12s. 6d.

This is one of the Foundations of Modern Psychology Series, a series in paperback form, covering the main branches of psychology in a set of small books, each written by an expert in his field, and prepared for the elementary student. Mussen admits in the preface that "no volume as brief as this could possibly present a complete or exhaustive survey of the principles, facts and theories of child development. . . Rather the book is designed to give the reader with little or no training in the field a broad perspective of its objectives, scope and contents." It succeeds admirably in fulfilling this aim. The book forms a clearly expressed, interesting survey of the main aspects of child development. The well-known experiments are all mentioned and their relevance clearly explained.

It is somewhat surprising that emotional development is not considered separately but merely touched upon under the headings of 'The Development of Personality' and 'Social Behaviour.' One must bear in mind, however, that there is a complete volume in the series devoted to 'Motivation and Emotion.' In these two chapters on personality and social development, some of the examples seem less applicable to this country than to the American culture from which they are drawn. This is, perhaps, an inevitable limitation of any American book on these aspects of development.

A criticism not only of this volume, but also of others in the series, concerns the lay-out of the text, which could be improved greatly and thereby provide maximum assistance to the elementary students for whom it is designed. Two examples from this volume must suffice. The table of contents is annoying—because this commences half-way down a page the reader is prevented from viewing the contents at a glance. Secondly, at the beginning of each chapter there is a page of double-spaced print which one assumes is an introduction. From Chapter 4 onwards, the double-spacing stops short at the end of the page as before, but in these later chapters this is not even the end of a sentence.

Child development, though a fascinating subject, is often found boring by elementary students. Reference to 'manuals' and 'handbooks' does little to relieve these feelings. This present volume will, therefore, be welcomed both by the students and lecturers dealing with this aspect of psychology. Its merit is further illustrated by the thought that there are those who may almost be tempted *not* to prescribe it because of the resulting loss of such a convenient scourge book for lecture material.

MARGARET M. CLARK.

OLIVER, R. A. C. (1964). *Studies in a University Entrance Test in English*. Manchester: Northern Universities Joint Matriculation Board, pp. 56, no price given.

The writer of this research pamphlet, in conjunction with the Joint Matriculation Board, has made some useful contributions to the science and art of examining both in English and in General Studies. Here we have yet another valuable contribution, which analyses statistically the results of several university entrance tests in English, from the first experimental trials to their official use in March and July, 1963. Though he finds that "nearly two-thirds of the candidates who twice failed

in the Test in English (March and July) gained two or more passes in the G.C.E. at the Advanced Level," he also concludes that the number "who would have failed to qualify for entrance to universities solely because of their failure on two occasions in the Test in English (had this requirement been compulsory) . . . is estimated to be less than 1 per cent. of the candidates who entered for the test in English in March, 1963." The wide difference between these figures is because a large percentage of those who failed twice in English but passed in two Advanced level subjects actually obtained only low grades in the latter, and it is conjectured that none of these would have been accepted by a university. This reviewer is very doubtful whether the rejection would have been quite so sweeping. None-the-less, it is true that the percentage rejected because of failure in English (*after two attempts*) would have been small, and that even if the rejections in 'A' level results had been less sweeping, the extra candidates admitted would on the average have been of low quality.

Another point of importance is that the low figures of 1 per cent. is achieved by using as a basis all those who entered for the test in English in March, 1963. Now this includes the candidates who proved to be unsuccessful in 'A' levels as well as in English. What we would have been much more interested in knowing is the percentage of those *successful in 'A' levels* who, nevertheless, failed twice in English. A relevant statistic here is that "some candidates who failed twice achieved high grades in the G.C.E. at the Advanced level or in Special papers or in both, e.g., 7.1 per cent. of them gained two B grades or better." Personally, the reviewer would be against excluding this 7.1 per cent. If a man can, for example, communicate at a high level in mathematics, should he be excluded because he fails in an examination in English?

On page 23 the author mentions in paragraph 30 that boys who took the Oxford and Cambridge paper had the incentive of securing a university entrance qualification, but there was no such incentive in the J.M.B. test. When, however, later on he attempts to analyse the reasons for so many more students failing the J.M.B. than failed the Oxford and Cambridge, he omits entirely to take this factor into account, while putting forward other possible reasons. There is a minor inexactitude on page 26 where it is stated that "at school N none of the individual questions agrees with the school estimates with the exception of J.M.B. Question 4"; in fact, four of the correlations are around 0.3, and though they are said in the table to be not statistically significant, this lack might well be due to the smallness of the number of candidates—though it is here admitted that 0.3 does not represent close agreement. In spite of these points of criticism, the pamphlet is well worthy of study and is a useful addition to work in this field.

R. R. DALE.

OLIVER, R. A. C. (1964). *The Marking of Scripts in Advanced Level History*. Manchester: Northern Universities Joint Matriculation Board, pp 36, no price given.

Despite the prestige enjoyed by the Boards which conduct G.C.E. examinations in Britain, they are open to criticism for their reluctance to publish the results of such investigations as they presumably do make of the reliability of their annual examinations. Wiseman (1961) voices the misgivings of many when he says: "What is disturbing, however, is the paucity of any published enquiries of this kind sponsored by the examining bodies themselves. This contrasts very markedly with the 11-plus selection, where a great deal of research has been carried out, or initiated, by those actually involved in the selection."

These misgivings are not allayed by the results of studies in other countries of comparable examinations. An investigation of the marker reliability for parts of the 'Bagrut' or Matriculation examination in Israel—a country with a standard of examining consonant with its high regard for academic excellence—showed unexpectedly low inter-marker correlations and led to procedural changes (Pilliner, 1961). Results obtained from an extensive study of an Indian University's

Matriculation examination were even more disconcerting (Taylor, 1964). Therefore, the publication earlier this year by the Joint Matriculation Board of the pamphlet OP17 is of special interest.

This pamphlet reports the results of a reliability study of marks awarded in the 1962 History (Advanced) examination set by one Board (Board 1). Photostat copies of the scripts of twenty candidates for Board 1's examination, selected to cover the range from 'Distinction' to 'Allowed Ordinary,' and already marked by two of this Board's examiners under the usual examination conditions, were marked again by examiners from six other Boards (Boards 2-7), after they had studied the first Board's syllabus, question papers and marking schemes. The comments and marks assigned by Board 1's examiners were retained on the scripts so as to be available to examiners from the other Boards who were asked to mark in two ways: firstly, in accordance with Board 1's marking scheme; and secondly, using their own procedures.

Board 1's marks or ranks were correlated in turn with the marks or ranks assigned by each of the other Boards. All six rank (and three product-moment) correlations are reported, ranging between .97-.99, with an arithmetic mean of .984. The pamphlet fails to make it clear whether the Board 2-7 marks or ranks concerned in these correlations are based on Board 1's marking scheme or on the schemes of the other Boards.

What do these correlations show? As stated in the pamphlet, the question asked by the experimenter was as follows: "Here is the work of twenty candidates who were arranged in an order of merit by examiners working under conditions normal for the Board which employs them. Now, some months later, and working under less pressure, what do the examiners of other Boards think of that order?" The high correlations reported show that under the conditions of the experiment the other examiners do indeed think very well of the original order.

But how valuable is this information, since the original marks and comments were left on the scripts for the other examiners to see? The close agreement recorded is wholly irrelevant to the marking situation which normally confronts an examining Board. "Only after considerable discussion," we are told, was the decision taken to leave the marks and comments on the scripts. However cogent the reasons for the purposes of the experimenters, that decision was surely the wrong one from the point of view of most readers of the pamphlet, who are likely to be more interested in the answer to a totally different question, that is: What correspondence is there between the marks or ranks assigned when two examiners, similarly qualified and briefed, *independently* mark the same set of scripts? The answer to this question, which relates to a realistic situation, would furnish intelligible and useful information about marker reliability. We need urgently to know what difference it makes to the marks and rank orders if a batch of scripts is sent to one examiner rather than another. The lack of awareness evinced in this pamphlet of the kind of investigation that is really needed is calculated to heighten misgivings rather than to allay them. Here was an opportunity to conduct an experiment the results of which might have dispelled some of the uncertainty that Wiseman deplores. The opportunity was missed. Instead, the experiment was so designed that the only statistical question that could be answered was one of minor importance. We know now, as we might have suspected from the start, that a number of examiners agreed about the verdict of two of them when the evidence on which that verdict was based was plainly displayed to all. We still do not know the answer to the question which most of us see as crucial.

A. E. G. PILLINER.

REFERENCES

- PILLINER, A. E. G. (1961). *The Measurement of Educational Achievement in Israel*. Report to UNESCO.
 TAYLOR, H. J. *Operation Passmark: An Account of the Methods Used in the Matriculation Examination of 1963*. Gauhati: University of Gauhati.
 WISEMAN, S. (Ed.) (1961). *Examinations and English Education*. Manchester: The University Press.

PER RAND (1963). *Distortion and Selectivity*. The Norwegian Research Council for Science and the Humanities, pp. xii, 272.

Like other Scandinavian doctoral theses, this book manifests a high degree of scholarship in the theoretical foundations of its empirical investigation. Over half the total space is devoted to a review of the literature. The author describes and evaluates results and ideas in a variety of fields concerned with the determinants of levels of aspiration and levels of performance in achievement situations. The particular determinants considered are selective and distorting factors in memory and perception, and anxiety. Each of these fields is covered with thoroughness; several infrequently mentioned, but significant details of well-known experiments enable Rand to make some interestingly original interpretations.

Rand uses three theoretical frameworks to integrate the literature; Postman's hypothesis theory, Festinger's cognitive dissonance theory and Atkinson's development of Lewin's analysis of level of aspiration. Controversial points in these theories are handled with clarity and confidence, but the final synthesis is difficult to grasp. Research workers in any of the areas of study mentioned will find Rand's account profitable reading.

A single complex investigation is used to explore the efficacy of the synthesized framework. Thirty-seven hypotheses are formulated and tested, but as with the synthesis, the derivation of hypotheses is difficult to follow: the hypotheses involve differential predictions for boys and girls differing in anxiety, but the links between sex differences and the previous discussion are not obvious. From the results Rand claims significant support for nineteen of the hypotheses. The measures used are checked for reliability (in fact, Rand is one of the few workers to have checked the internal validity of the Sarason General Anxiety Scale), but the investigator is willing to accept a rejection of the null hypotheses at the 10 per cent. level with one tail tests. If a 5 per cent. two tail criterion is imposed, only nine hypotheses are confirmed. There is only a short discussion of the reasons why the elegantly derived theoretical principles were not more strongly supported. The absence of prolonged critical discussion is to be regretted, the more so because the author has exhibited considerable critical skill in the first half of the book.

RIVERS, WILGA M. (1964). *The Psychologist and the Foreign-Language Teacher*. The University of Chicago Press, pp. 212, \$4.00.

The direct contribution of psychologists to the methodology of teaching specific subjects at the secondary school level has, at least up to very recently, been extremely limited, and it is unlikely that many practising teachers or students in training even attempt to bridge the gap between psychological theories of learning and the problems of learning involved in the subjects with which they are concerned. This book, therefore, is to be welcomed as a thoughtful and interesting attempt at linking the contributions of the major modern learning theorists with the teaching of foreign languages, with particular reference to French.

As a teacher of languages the author is clearly in sympathy with audio-lingual methods—or oral/aural techniques as they are more often termed in this country. Recently, the growing study of linguistics has helped to give audio-lingual methodology a more uncompromising body of theory, the most notable outcome of which has been a flourishing of language laboratory techniques. In this book Dr. Rivers, whose concise style is marred only occasionally by a tendency to repetition, puts these recent developments into perspective by examining the support which they can find in modern learning theory. Some readers may regret that she has confined herself to an examination of audio-lingual methodology, thereby limiting her treatment of the place of visual stimuli in language learning. However, the issues which Dr. Rivers examines are the ones which are uppermost in the minds of language teachers at the moment, in particular, the question of the effectiveness of teaching techniques based on oral repetition and analogy rather than translation and grammatical analysis.

Dr. Rivers certainly knows her way about the mazes of learning theory and she is scrupulous in arguing her methods from psychological principles. She rightly stresses the importance of communication in language learning and suggests that mechanical language laboratory techniques, valuable as they may be, are limited in their effectiveness and "must be supplemented by communication experiences where desire to be understood is satisfied." At a time when machines are increasingly influencing audio-lingual methodology, she uses the theories of the Neo-Behaviourists to underline the importance of the personal role of the language teacher. Not merely must he provide the right models for imitation; he alone can help the learner to avoid fatigue and boredom and can create the best emotional climate for learning to take place.

This book is aimed at practising teachers and deserves to be read carefully by them. It is not difficult to read though the reader has to remain mentally alert throughout. As, however, consideration of practical techniques is subordinated to consideration of theory, it is likely to be most appreciated by teachers with a more than average amount of psychological sophistication. It will certainly interest all, including students, who are concerned with the methodology of language teaching in university departments of education and training colleges.

Educational psychologists who would like to make a contribution to this field will also benefit from studying this book, which is well documented and contains a large number of references to work on linguistics, language teaching and learning theory. The appendix, a summary of the work of the major learning theorists, is well written, and is followed by a useful bibliography.

MAURICE CHAZAN and HAROLD ROTHERA.

ROUCEK, JOSEPH (Edit.) (1964). *The Difficult Child*. London: Peter Owen, pp. viii+292, 35s.

This is a collection of sixteen contributions from various authors. One contribution is on poor memory, one on family counselling, while each of the remaining fourteen deals with a different kind of difficult child. A short selected bibliography and a thumbnail sketch of the contributor usually follows each set of references (the references are called footnotes). The authors are American and the bibliography and references are overwhelmingly concerned with American material. The book is intended for teachers, social workers and parents.

The term 'difficult child' is interpreted somewhat unusually. While it may be reasonable to find both the aggressive child and the poor reader included, the creative child and the precocious child seem out of place in this collection. Admittedly, Torrance is at pains to argue his thesis that the creative child is a difficult child, but your reviewer is not convinced. The pampered child is in, but not the withdrawn child nor the nervous child, two of several categories of difficult child which might be considered more easily recognisable entities. Some of these anomalies may be explained by reference to the transatlantic culture for which this book was written, but even allowing for this the selection of topics still seems arbitrary.

Inevitably, attempts to deal with such broad topics in eighteen pages or so will be superficial, as Barnes, writing on the poor reader, recognises. Inevitably, too, one would expect that topics which are hard to define will lead to contributions which overlap. Thus, the contribution on the precocious child contains a section on the creative child and that on the amathematical child contains a section on the inattentive child.

The level of presentation varies considerably. One or two topics are dealt with in a way which would be useful introductory reading for students. Others are written at a frankly popular level, well garnished with anecdotes. Clearly both approaches have their place, but their places are different.

The selected bibliographies give brief outlines of the tests recommended. This is a helpful practice. But, in Scherer's contribution on the inattentive child, the bibliography seems really to be an alphabetically arranged list of references. Other errors make the referencing of this particular contribution especially difficult to follow.

Recent work on children with behaviour problems has revealed many targets for text books. This one aims at several, but cannot be said to have scored any direct hits.

PHILLIP WILLIAMS.

VERNON, M. D. (1962). *The Psychology of Perception*. Harmondsworth: Penguin Books, pp. 264, 5s.

The high standard achieved by the Pelican Psychology Series is maintained. Readers of this *Journal* are likely to be interested in this addition to the series: it is particularly strong on the developmental side. As Professor Mace says in his editorial Foreword, the book keeps in mind the interests of parents and teachers as well as those of the general reader.

The author has included a great deal of information: Gestalt psychology, as represented also by Gottschaldt and Loretta Bender; Ames and Michotte; Thurstone's factorial study; and relations between Perception and Personality. Many classical experiments are reported: Sherif on the autokinetic effect; Bruner and Goodman on size perception; Murphy and Schafer on Punishment; Carmichael, *et al.*, on effects of naming. Appropriately, critical evaluation is given to some of the earlier 'functional perception' researches. Of particular interest to psychologists and teachers is the Chapter on 'Perception of Special Types of Material.' This includes an account of Professor Vernon's own studies of visual aids including diagrams and graphs.

Inevitably, the reviewer finds omissions. On the whole, emphasis is placed on exteroception in general, and vision in particular, though one chapter deals with relations of perception to motivation and emotion. The perception of time is omitted by the author, as by many before her. This seems a pity as, in the work of Frankenheuser, Loehlin, Fraisse and others, advances are being made in this difficult field. Another omission is the work of Gregory on visual illusions. Mental imagery is rather briefly dealt with. This is to be regretted at a time both when American engineering psychology (e.g., high altitude flying, polar exploration, and motorway driving) is bringing this neglected subject again to the fore, and when the Dement-Kleitman break through with EEG. and rapid eye movement procedures has produced techniques for the more rigorous study of waking as well as dream imagery. The related subject of synaesthesia—rather oddly described as "one curious peculiarity of colour vision" (p. 83)—is somewhat briefly dealt with, particularly in the light of Professor Vernon's own interesting treatment of this in an earlier book.

Chapters on 'How the Child Learns to Perceive,' and on 'Development of Shapes and Colour Perception in Children,' will be of special interest to readers of this *Journal*. As Professor Mace points out in his Foreword, this book is an attempt to formulate in a clear way, questions that can be answered by scientific procedures. It seeks to provide such answers as exist at the present state of knowledge. The author is to be congratulated on her achievement of these aims in this lucidly written book.

PETER MCKELLAR.

WEPMAN, J. M., and HEINE, R. W. (Edit.) (1964). *Concepts of Personality*. London: Methuen and Co., Ltd., xxxix+514, 70s.

Although the literature on personality studies is now voluminous, there remains a shortage of books of high quality, particularly those covering the whole field. Hall and Linzey's *Theories of Personality* is an outstandingly good book of this kind, and the present volume attempts much the same task. It covers approaches to the subject by means of learning theory, perception, genetics, drive theory, field theory, psychoanalysis, humanism, personal constructs, behaviorism, phenomenology, sociology, anthropology, factor analysis and clinical assessment, and the authors include Cattell, Fiske, Eriksen, Kounin, and other well-known writers. It is more up-to-date than Hall and Linzey, but written at a rather more elementary level. Nevertheless, it is to be recommended to both the specialist and general reader.

F. W. WARBURTON.

PUBLICATIONS RECEIVED

The mention of a book in this list neither implies, nor precludes, a later review.

- J. D. HALLORAN (1964). *The Effects of Mass Communication*. Leicester University Press, pp. 83, 7s. 6d.
- STRANGE ROSS (1964). *Logical Foundations of Psychological Measurement*. Copenhagen: Munksgaard, pp. 152, 3 kr.
- JOSEPH ROUCEK (Edit.) (1964). *The Difficult Child*. London: Peter Owen, Ltd., pp. viii+292, 35s.
- FRED N. KERLINGER (1964). *Foundations of Behavioral Research*. New York: Holt, Rinehart and Winston, pp. xix+739, no price given.
- VIKTOR LOWENFELD and W. LAMBERT BRITAIN (1964). *Creative and Mental Growth*. New York: Macmillan Co., pp. xv+412, 55s., fourth edition.
- WILLIAM K. DURR (1964). *The Gifted Student*. New York: Oxford University Press, pp. vi+296, 44s.
- UNIVERSITY OF HARVARD (1964). *List of Books in Psychology*. Harvard University Press; London: Oxford University Press, pp. viii+111, 20s., third edition.
- SIR FREDERIC BARTLETT (1964). *Thinking*. London: Unwin University Books, pp. 203, 15s., third edition.
- LARS H. ALIN (1964). *Experimental Studies in Verbal Versus Figural Learning and Retention*. Stockholm: Almqvist and Wiksell, pp. 216, Swedish kronor 40.
- RAY HYMAN (1964). *The Nature of Psychological Inquiry*. New Jersey: Prentice-Hall, pp. xi+116, 12s. 6d.
- EDWARD J. MURRAY (1964). *Motivation and Emotion*. New Jersey: Prentice-Hall, pp. ix+118, 12s. 6d.
- JOHN B. CARROLL (1964). *Language and Thought*. New Jersey: Prentice Hall, pp. x+118, 12s. 6d.
- R. J. HAVIGHURST, MARIA EUGENIA DUBOIS, M. CSIKSZENTMIHALYI and R. DOLL (1964). *A Cross-National Study of Buenos Aires and Chicago Adolescents*. Basel: S. Karger, A.G., pp. 80, sFR/Dm 15.
- ARTHUR W. STAATS (1964). *Human Learning*. New York: Holt, Rinehart and Winston, pp. xv+520, \$8.75.
- D. E. BROADBENT (1964). *Behaviour*. London: Methuen and Co., Ltd., pp. 215, no price given. First published in 1961 by Eyre and Spottiswoode (Publishers), Ltd.
- RONALD A. KINCHLA (1964). *Statistics*. Reading, Mass., and London: Addison-Wesley Publishing Co., pp. 478, 40s.
- H. OLDMAN (1964). *Educational Television*. Leeds: University of Leeds Institute of Education, pp. 46, no price given.
- MARIA MONTESSORI (1964). *Spontaneous Activity in Education*. Cambridge, Mass.: Robert Bentley, Inc., pp. xi+355, \$6.50.
- MARIA MONTESSORI (1964). *The Montessori Elementary Material*. Cambridge, Mass.: Robert Bentley, Inc., pp. xviii+464, \$8.50.
- GEORGIA SACHS ADAMS (1964). *Measurement and Evaluation in Education, Psychology and Guidance*. New York: Holt, Rinehart and Winston, pp. xiii+654, \$8.95.
- LESLIE T. WILKINS (1964). *Social Deviance*. London: Tavistock Publications, pp. xv+311, 45s.
- CLYDE H. COOMBS (1964). *A Theory of Data*. London: John Wiley and Sons, Ltd. pp. xviii+585, £5 13s.
- KATHLEEN GIBBERED (Edit.) (1964). *Your Teenage Children*. (London: Macdonald pp. 86, 10s. 6d.
- WALTER B. BARBE (1963). *The Exceptional Child*. Washington, D.C.: The Center for Applied Research in Education, Inc., pp. ix+117, 32s.
- RALPH GARRY (1963). *The Psychology of Learning*. Washington, D.C.: The Center for Applied Research in Education, Inc., pp. viii+111, 32s.

- ROBERT J. C. HARPER, CHARLES C. ANDERSON, CLIFFORD M. CHRISTENSEN and STEVEN M. HUNKA (1964). *The Cognitive Processes, Readings*. London: Prentice-Hall International, pp. xii+717, 72s.
- JOHN CURTIS GOWAN and GEORGE D. DEMOS (1964). *The Education and Guidance of the Ablest*. Illinois: Charles C. Thomas, pp. xi+511, \$14.50.
- G. W. PARKYN (1964). *The Mental Health of the Gifted Child*. Wellington: New Zealand Council for Educational Research, pp. 24, 1s. 6d.
- BRIAN SIMON (Edit.) (1964). *Non-Streaming in the Junior School*. Leicester: P.W.S. (Educational) Publications, pp. 90, 5s. 5d.
- P. A. E. GRADY and J. C. DANIELS (1964). *A Survey of the Incidence of Speech Defect in Children*. University of Nottingham Institute of Education, pp. 20, 2s. 6d.
- BARRY E. COLLINS and HAROLD GUETZKOW (1964). *A Social Psychology of Group Processes for Decision-Making*. New York: John Wiley and Sons, pp. x+254, 47s.

ENVIRONMENTAL HANDICAPS AND INTELLECTUAL
DEVELOPMENT: PART II

By PHILIP E. VERNON

(University of London, Institute of Education)

SUMMARY. This article gives the comparative results of samples of West Indian and English 10½-11-year boys on a battery of perceptual and performance tests. These, together with the results on educational and verbal tests and assessments of background variables, were analysed, and a similar factor structure was obtained in the two groups.

The deficit of West Indian boys varies considerably with different tests, being most marked for practical and some non-verbal *g* tests (also for pure linguistic development), much less serious for educational attainments and for various perceptual abilities. Differences between urban and rural sub-groups in the West Indian sample tend to replicate this pattern.

Substantial correlations in both samples were obtained between test performance and assessments of the cultural stimulus provided by the home, and smaller ones with socio-economic status and with purposeful-rational home climate. Male influence in the home seems to be particularly associated with the development of more practical abilities. The implications of the research for testing the educational potential of West Indian immigrant children in Britain are discussed.

IV.—THE NON-VERBAL TESTS AND THEIR RESULTS.

Matrices Test. A twenty-four-item Matrices test, in which the testee draws his own solutions, was constructed and administered individually, after having been tried out and item-analyzed in an English primary school. In each of the first eleven items, the testee was asked to give the reason for his response and, if he was wrong, the principle was explained and he was allowed to correct the response (though scored wrong). It was considered that this was a more satisfactory method of training on the test than any system of group demonstration as advocated by Schwarz (1961), MacArthur, *et al.* (1964) and others.

The median West Indian performance was equivalent to an English quotient of 75 (see Table 3), according to the method described in Part I of the article.

Porteus Mazes. In the English group, the standard series of mazes from V to XIV was applied to alternate boys by one of two methods: the first method was that specified by the author, where a fresh maze is given as soon as an error is made. By the second method, the testee is allowed only one copy of each maze and goes on drawing after correcting any error. The two methods yielded almost identical distributions of Maze Mental Ages. Thus, the second one is to be preferred as being much more economical of printed blanks. Note that the present-day performance of English boys is much superior to Porteus's norms, the median M.A. for 11.0 year-olds being 13. Possibly this is due to their familiarity with mazes published in Comics. The West Indian group (all tested by the second method) also scored above the norms, with a median M.A. of 12, though they would have much less opportunity for previous familiarization. Compared with our English group, the median quotient is 91.

Kohs Blocks. The WISC Designs were used, but with certain modifications in procedure suggested by Goldstein and Scheerer (1941) and developed by McConnell (1954) and Jahoda (1956). Models A and B and Card 1 were given as usual, but if the Subject failed on Card 1 or any later card:

(a) The tester demonstrated the model, destroyed it and asked S to try again.

(b) The tester made the model and left it for S to copy.

(c) The tester made it once again and if S still failed on two successive cards, the test was abandoned.

A score of 0 was given if (b) or (c) was needed, 1 for (a), and higher scores (similar to Wechsler's) if the design was reproduced without assistance in various

TABLE 3

COMPARATIVE PERFORMANCE OF ENGLISH AND WEST INDIAN BOYS ON NON-VERBAL TESTS.

Test	English Percentiles			West Indian Percentiles			West Indian Deviation Quotients		
	90	50	10	90	50	10	90	50	10
Matrices	21½	19	8	17	6½	3	93	75	—65
Porteus Mazes ...	15½	13	10½	14	12	8	107	91	65
Kohs Blocks	39	26	13	23	7	1	96	75	—65
Vernon Form-board	31	25	20	23	16	8	92	68	—65
Embedded Figures	18—	14—	10½	15½	12.	6½	104	88	—65
Reproducing Designs	6	11½	17	7½	14½	18½	113	87	76
Draw-a-Man	30½	24½	17½	32½	21	12	128	91	—65

lengths of time. In the West Indian sample, some 40 per cent. were virtually unable to transfer from copying models to copying printed designs, despite demonstration (a); that is, their total score was 4 or less. Only 4 per cent. of English boys did as badly, though several showed unusual blockages or rotations (cf. Shapiro, 1954) at a rather higher level of performance. The median West Indian quotient was 75.

Vernon Formboard. This consists of seven wooden pieces and seven different matrices into each of which all the pieces can be fitted. The median times for solving successive matrices range from fifteen seconds to about two minutes, and scores from 0 to 5 are given for various speeds of completion of each matrix. If a Subject works unsuccessfully at filling any one hole in a matrix for over thirty seconds, he can be shown the location of a key piece, and thirty seconds is added to his total time for that matrix. Admittedly, this introduces a certain chance element into his score, but as the final score is based on seven different tasks, it is likely to be more reliable than that obtained from formboards containing only a single set of holes.

The median West Indian performance was lower on this test than any other, corresponding to a quotient of 68, thus contradicting any easy generalization that backwardness arises primarily from linguistic or educational handicap. Over 50 per cent. of the boys were as slow as, or slower than, the bottom two boys in the English sample.

Three-dimensional Perception. In the West Indies, six black and white drawings involving 3-D perception were presented, for example, a nearby aeroplane and a distant airport, a boy walking along a road and a man further back, a chain of mountains in a valley; and the Subject was questioned as to which was the biggest.

With the English sample, the test was modified and made up to ten drawings by including some of the rather difficult 3-D pictures devised by Hudson (1960) and applied in Sierra Leone by Dawson (1963). Scores on these two versions are obviously not comparable. However, 60 per cent. of the West Indian sample showed definite errors of perception or recognition on relatively easy pictures, while 31 per cent. of English boys made from three to five errors on the more difficult series.

Embedded (or Gottschaldt) Figures. A twenty-one-item test was constructed and item-analysed on the basis of trials in English schools. In each item a simple figure is shown on the left; the Subject has to perceive this within a more complex figure on the right, and trace it out with red pencil. In the West Indies the test was given individually with ample explanation and assistance, but it was found that blackboard demonstration to each group of boys was adequate in England. There was no time limit. It is possible, though unlikely, that the difference in procedure helped the West Indies sample, whose median quotient was 85.

Reproducing Designs. This and the Draw-a-Man test were likewise given in group form in England, but individually to West Indian boys in order to ensure full understanding of what was required. Four items (A, 4, 5 and 6) were taken from the Bender-Gestalt, and combined with the Memory for Designs items from Terman-Merrill (M.IX.1 and L.IX.3). In the group version the demonstration drawings were magnified to about six times the original size.

For each figure, four to six common errors were listed and carefully defined, and each Subject was scored for the total of such errors. The median West Indian score corresponded to an English quotient of 87. Here, too, they were at some advantage since, with individual administration, the tester can more readily get the Subject to concentrate on the task.

Draw-a-Man. The standard Goodenough instructions and scoring were used. The English median of $24\frac{1}{2}$ points was very low, corresponding to a Mental Age of 9 years; but this commonly happens with older junior pupils. Possibly, the West Indians were rather better motivated, and their median of 21—though little over 8 years by Goodenough's norms—falls at an English quotient of 91.

Mischel's Chocolate Test. No attempt was made to test social or educational attitudes; but largely with a view to rewarding the boys for their co-operation, a test developed by Mischel (1958) was included. After the first individual testing session each boy was offered the choice of a small bar of chocolate now, or a double-sized bar when all the testing was completed a week or so later. Working in the West Indies, Mischel presents some evidence that this test of willingness to delay gratification is associated with a more secure family background and with masculine salience in the home. 75 per cent. of the English and 54 per cent. of the West Indian boys chose the large bar later.

V.—ANALYSIS OF TEST RESULTS AND ASSESSMENTS.

Factor Analyses. We have two main types of results to consider—differences in test performance between cultural groups which we hope to link up with

observed differences in child upbringing and schooling (cf. Vernon, 1962), and individual differences within groups which can be correlated with the assessments of environmental factors. In addition, differences between the contrasted West Indian sub-cultures may be enlightening. We would hope that these three approaches will complement one another. However, it would be unwise to base any interpretations on subjective notions of what abilities the tests measure; and speculation is particularly hazardous within an unfamiliar cultural group where, as Irvine (1964) has pointed out, test results may be affected by numerous influences that do not operate in the familiar culture. By applying factor analysis we can tread more firmly, since we can both see whether the tests group together in the same fashion in the two cultures, and can more readily identify the abilities underlying such groups of tests.

Several techniques of analysis were applied to both sets of scores, and Principal Components* for the English figures, together with a Varimax rotation, were kindly supplied by Mr. H. J. Hallworth, of the Birmingham University Department of Education. After some further adjustments to the rotations, the solution shown in Table 4 was obtained.

It will be seen that, as usual, there is a large general factor running through all the tests with its highest loadings on Arithmetic, Abstraction, Matrices and Kohs. An Educational factor loads all the verbal tests, and there is no differentiation of a memorizing nor an arithmetic factor. But some additional variance is present among the two Vocabulary and English tests and Concept Formation, hence a further Verbal factor is distinguished. This is quite small, hence it is not surprising that some of the loadings appear illogical, particularly the positive loading for Draw-a-Man and the negative one for Porteus Mazes.

The conventional spatial dimension likewise splits into two. The larger one, labelled Perceptual, particularly loads Embedded Figures, Reproducing Designs, Draw-a-Man and Picture Recognition, but also enters into Piaget Visualization and into Word Memory, presumably because the latter tends to involve some visualization of word lists. The smaller component, labelled Practical, is mainly found in Piaget Visualization and Formboard, with non-significant loadings for Kohs Blocks and Draw-a-Man. The sixth factor loads all the interview assessments, whereas all the test-loadings are non-significant; thus, it may be identified as a kind of halo effect of 'the good home.' We will comment later on the correlations of environmental variables with the test factors.

The Centroid technique was applied to the test correlations in the West Indian sample and four factors extracted, the loadings of the environmental variables on these being calculated subsequently. The general pattern was the same as in the English sample, and the variances of the general and group factors closely comparable. The main difference appeared to be that three tests—Matrices, Concept Formation and Piaget CV gave the purest measure of *g*, almost all the other tests showing an additional common factor which could be interpreted, partly as schooling, partly as familiarity in dealing with pictures and symbols as well as with words. The three tests just mentioned appear to get most naturally and directly at intelligence in the sense of grasping relations, and to be least affected by scholastic experience. In view of the smallness of the sample and the consequent unreliability of the correlations, it was difficult to achieve a satisfactory rotational solution of the Centroid analysis. Instead, therefore, a group factor analysis is quoted in Table 5.

* The communalities for insertion in the diagonals were estimated from a preliminary Centroid analysis of the test inter-correlations.

TABLE 4

ROTATED PRINCIPAL COMPONENTS OF VARIABLES IN ENGLISH SAMPLE.

Test or Other Variable	Factor Loadings						h ²
	I 'g'	II Educ.	III Verb.	IV Perc.	V Prac.	VI Halo	
Arithmetic	.71	.52	-.01	.04	.06	.01	.78
English	.62	.65	.22	.09	-.03	-.03	.86
Vocabulary A (Group)	.61	.55	.35	.14	.07	.05	.83
Vocabulary B (Individual)	.56	.33	.43	.11	-.06	.15	.65
Word Memory	.51	.46	.06	.20	.11	-.13	.54
Information Memory	.62	.50	-.05	-.11	-.01	.14	.67
Abstraction	.72	.51	.07	-.09	.04	.00	.79
Piaget Arithmetical	.71	.04	.01	-.08	.04	.17	.55
Piaget Conservation	.68	.09	-.13	.14	.14	.03	.52
Piaget Visualization	.48	-.07	-.13	.34	.39	.12	.54
Matrices	.79	.08	-.02	.10	.02	-.11	.64
Concept Formation	.54	-.12	.31	.17	-.09	-.01	.44
Porteus Mazes	.50	.13	-.26	.05	-.12	-.07	.35
Picture Recognition	.39	-.11	.01	.31	-.13	-.13	.29
Embedded Figures	.68	.09	.08	.42	-.09	.04	.66
Reproducing Designs	.61	.10	-.04	.41	.01	-.06	.56
Kohs Blocks	.80	-.09	-.04	.22	.10	.11	.73
Draw-a-Man	.43	-.14	.21	.50	.18	-.10	.54
Formboard	.53	.12	-.09	.23	.31	-.17	.48
Percentage Variance	28.7 37.8	8.7 10.2	2.9 3.3	4.4 5.7	2.3 2.0	4.0 1.0	— 60.0
Delayed Gratification Test	.07	.20	-.15	.27	-.21	.21	.23
Regularity of Schooling	-.15	.31	.07	.21	.05	.13	.18
Unbroken Home	-.11	.30	.01	.09	.14	.43	.31
Socio-Economic Status	.38	.18	.24	3.0	.10	.37	.37
Cultural Stimulus	.56	.29	.16	.01	.07	.38	.56
Initiative	.13	.03	.00	.13	.03	.31	.13
Male Dominance	-.06	.05	.02	.30	.39	.12	.17
Planfulness	.32	.26	-.03	.03	.11	.50	.43
Linguistic Background	.49	.31	.24	.04	-.10	.28	.48

Loadings of psychological interest are shown in heavy type. Those of about .20 up are statistically significant.

Here again there is no trace of a memorizing factor, though a separate verbal factor is suggested (as in the English group) by the negative *h*-loading for Vocabulary B. Only one spatial factor seemed to be needed. However, the Centroid analysis indicated a possible differentiation between Picture Recognition+Design Reproduction on the one hand, and the more practical Porteus Mazes and Formboard on the other. (In the English group, Mazes showed no spatial content). The Piaget CV items, as mentioned, act as a better *g*-measure here, whereas Piaget A (arithmetical-orientational) items show more spatial content.

Cultural Differences. We may now consider the relative strengths and weaknesses of the West Indian boys on different tests (shown in Tables 1 and 3) in relation to the factorial structure.

TABLE 5

GROUP FACTOR ANALYSIS OF VARIABLES IN WEST INDIAN SAMPLE.

Test or Other Variable	<i>g</i>	<i>v : ed</i>	<i>k</i>	<i>h</i> ²
Arithmetic69	.59		.82
English59	.71		.85
Vocabulary A64	.57		.74
Vocabulary B61	.15	— .18	.43
Word Memory55	.69		.78
Information Memory69	.59		.83
Piaget Arithmetical67	.30	.24	.59
Piaget Cons. Visual65			.42
Matrices*95			.91
Concept Formation50			.24
Porteus Mazes60	.18		.39
Picture Recognition55	.17		.33
Embedded Figures61		.43	.57
Reproducing Designs47	.45	.60	.78
Kohs Blocks74	.18	.35	.71
Draw-a-Man57	.25	.61	.76
Formboard57		.24	.40
Percentage Variance	40.4	14.7	6.9	62.1
Delayed Gratification	— .07	— .23	— .35	
Length and Regularity of Schooling23	.19	.26	
Unbroken Home	— .17	.24	.15	
Stable home background14	.17	.02	
Socio-Economic Status27	.28	— .05	
Cultural Stimulus33	.46	.25	
Initiative27	.16	.03	
Male Dominance25	.03	.27	
Planfulness32	.07	.03	
Linguistic Background47	.41	.28	

* The excessively high *g*-loading for Matrices results from basing the *g*-factor on three tests only, instead of on the whole battery.

In the *v : ed* group of abilities it is clear that the West Indians do well on formal or rote attainments such as Spelling, Arithmetic and Word Memory; they are intermediate on Vocabulary A and English and do very poorly on more purely linguistic tests—Vocabulary B and Information Memory. This does not quite correspond to the factorial distinction between our Education and Verbal factors, though it resembles it.

Similarly, there is some, though incomplete, correspondence among the spatial abilities. The West Indians do relatively well on the primarily pictorial tests entering into our Perceptual factor—Draw-a-Man, Embedded Figures, Design Reproduction and probably Picture Recognition. They tend to be weakest on more practical abilities, notably Formboard and Kohs Blocks and some of the Piaget Conservation and Visualization items.

Generalization is most difficult in the case of tests which could be regarded as relatively pure measures of *g*—Matrices, Abstraction,† Concept Formation and possibly some of the Piaget items and Porteus Mazes. The Matrices score was very low; Abstraction was not given on this occasion but was certainly weak in an earlier study (Vernon, 1961); Concept Formation and Mazes were

† In the present analysis (Table 4), Abstraction shows an unusually high Educational loading.

quite good. This inconsistency is perhaps to be expected. The stronger and clearer the group-factor content of a test in a Western culture, the more does it seem likely to measure comparable abilities in other cultures; whereas relatively pure or alleged culture-fair tests may be the most likely to differ in what they measure from one culture to another, and the most susceptible to uncontrolled influences of upbringing, education and test-administration.

Sub-cultural Differences. Analysis of variance was applied to the differences between the means of the five West Indian schools. Although each school was represented by ten boys only, between-group differences were significant at the .01 or .05 level on the majority of tests, and there was a general tendency, with some exceptions, for the F-ratios to parallel the major cultural differences. That is, when the whole West Indian sample falls much behind the English sample, the West Indian rural boys also tend to fall most below the urban ones.

The biggest sub-cultural differences (.01) occur in:

Matrices, Kohs, Design Reproduction and Draw-a-Man, i.e. on the predominantly *g* and perceptual tests. (Design Reproduction and Draw-a-Man were, however, fairly good in the sample as a whole).

Intermediate differences (.05) occur in:

Embedded Figures, Picture Recognition, English, Word Memory, Vocabulary A, Formboard.

Non-significant differences occur in:

Arithmetic, Spelling, Information Memory, Piaget, Porteus Mazes, and Vocabulary B, i.e., in the predominantly *v:ed* tests. (Note, however, that Vocabulary B and Information Memory yielded extremely low overall means).

It would seem to follow that the major variations within the culture are not attributable so much to variations in the quality of schooling, which would particularly affect the third group of tests, as to other environmental factors.

On most tests, the school medians descended in the same order from good urban to remote small-holding. However, the 'slum' school did particularly well on Porteus Mazes, Gottschaldt and Formboard, particularly badly on Vocabulary A and B, English, Concept Formation and Word Memory. The pattern is similar to that sometimes found with working and middle-class populations in England (cf. Bernstein, 1961).

Correlations with Environmental Variables. In both groups the assessment of Cultural Stimulus in the home gives much the highest correlations—mainly with *g* and verbal abilities, while the Socio-Economic variable gives similar but rather lower figures. Note that Cultural Stimulus (*not* Economic Status) appears to have stimulated even the spatial abilities in the West Indies, presumably because less educated parents there tend to discourage constructive play as well as failing to develop verbal concepts.

The rating of Linguistic Background gives much the same pattern of loadings as Cultural Stimulus, but it is doubtful how far the teachers and interviewers were really assessing speech in the child's family and local environment; they may have been more influenced by the child's present attainment. In other words, the correlations may reflect the result of poor upbringing and schooling rather than a cause of intellectual retardation.

Family Planfulness and Rationality seems to be particularly relevant to *g* development, showing a small Education loading only in the English sample.

Particularly surprising is the negative correlation of Unbroken Home with *g* in both samples. Remembering that we are not concerned here with school attainment, it is psychologically conceivable that the child who lacks the security of stable parents and a father may develop a more questioning, insightful mind. Why Regular Schooling should also show a (non-significant) negative *g*-loading in England, though correlating positively with Educational and Perceptual abilities, is even more obscure. In the West Indian group it correlates with all types of ability suggesting that, in cultures where home environments may be highly deprived, attendance even at poor schools may stimulate mental development all round.

The rating for Initiative and Maturity in Play and Household Activities was disappointing, perhaps because its significance was difficult for the raters to grasp. It has no significant loadings in the English group, except for Halo. And, among the West Indians, it correlates better with *g* and *v*: *ed* than with *k* abilities, as Witkin's (1962) work would have led us to expect. However, in both groups, Male Dominance obtains the highest loading on spatial abilities (in the English group on the practical, not the perceptual, sub-factor). It also contributes to *g* in the West Indian, but not the English, sample.

Delayed Gratification Test. This showed small, but nevertheless often consistent, correlations with ability tests and environmental variables in the two groups.

TABLE 6
CORRELATIONS AND FACTOR LOADINGS OF DELAYED GRATIFICATION TEST.

Environmental Variable	English	West Indian	Test Factor	English	West Indian
Schooling141	.138	<i>g</i>07	-.07
Unbroken Home241	.229	Educational20	
Stable Background ..		.326	Verbal	-.15	-.23
Socio-Economic	-.083	-.146	Perceptual27	
Cultural Stimulus135	-.114	Practical	-.21	-.35
Initiative213	.020	Halo21	
Male Dominance	-.165	.006			
Planfulness134	-.171			

In both groups it is more strongly related to Unbroken Home and Initiative (and especially to Stable Background in the West Indies) than to Socio-Economic, Cultural Level, Male Dominance or Planfulness. Thus, there is some confirmation for Mischel's claim that the test reflects the influence of a stable background, but none at all that it goes with masculine influence nor purposefulness—a variable which should surely involve willingness to delay gratification.

Some further light was thrown by the West Indian school differences. Unexpectedly, the greatest delayed gratification (seven or eight out of ten boys choosing the large bar of chocolate later) occurred in the slum and small-town schools, the least delay (only two out of ten) in the better-class urban school. Both rural schools yielded five out of ten. This would suggest that chocolate is fairly commonplace to middle and upper working class boys, hence, they might as well enjoy the small bar now, whereas among the poorer boys a large bar is such a rarity as to be worth waiting for. In other words, the results may depend more on rather superficial social attitudes than on deep-seated Ego and Superego trends.

Under these conditions, it is not surprising that the correlations with ability factors, though substantial, are discrepant in the two cultures. However, it is interesting that in both there was a positive specific correlation with Porteus Mazes—often claimed as a test of planning capacity.

There were many other suggestive specific residual correlations between tests and environmental variables, but as these were mostly on the borderline of statistical significance, it would be wiser to collect more data before trying to interpret them.

VI.—GENERAL CONCLUSIONS.

The major finding of the study so far is that children whose mental development is handicapped by poor socio-economic, cultural and linguistic environment, by defective education and family instability, show this to a greater extent in practical-spatial and some abstract non-verbal abilities than they do in actual educational achievements. This is borne out by the urban-rural sub-group differences within the West Indies. Even an educational system which is grossly affected by lack of funds, buildings and equipment, brief and irregular attendance, poor quality of teaching and over-formality, produces fairly good attainments, especially in the more mechanical subjects, though there is no indication that it trains a general rote-learning ability. Probably it also stimulates development in non-educational abilities when the home environment is very deprived. In the English situation, differences in regularity of schooling seem to have much less widespread effects. It must not be forgotten, however, that more able children from better homes are naturally more apt to attend school regularly, hence the causal effects of schooling as such are difficult to disentangle.

Next it is clear that the most important single factor in children's performance on *g* and verbal tests is the cultural level of the home, parental education and encouragement, reading facilities and probably the speech background (though it was not possible to assess the latter aspect satisfactorily). Economic level as such is subsidiary, and among West Indian (not English) children, Cultural Stimulus strongly affects perceptual-practical abilities as well. Of the other factors in the home situation, suggested as relevant in an earlier publication (Vernon 1962), the most important seems to be the purposefulness and rationality of the family atmosphere. The broken home syndrome and family instability are to some extent associated with educational achievement but not at all, or even inversely, with the general ability factor. A definite connection was found in both groups between male dominance in the home or masculine identification on the part of the boy, and certain more practical abilities. Absence of a father figure is, indeed, an all too frequent characteristic of the West Indian culture.

Many psychologists and anthropologists are justifiably suspicious of the application of British or American tests in non-Western cultures, on the grounds that they may reflect different sources of variance, different background factors, effects of administration procedures, etc. In fact, the overall factor structure was extremely similar in the two groups, although certain tests differed sufficiently in their loadings to suggest that they were measuring different things. We are entitled to generalise that West Indian boys are relatively high in Educational and Perceptual, lower in *g* and linguistic, and lowest in the Practical factor; but must, at the same time, admit that there were considerable inconsistencies in their performance on some of the tests that go to make up these factors. Particularly, was this true of the tests which, in

Western cultures, may be regarded as relatively pure measures of *g*. Thus, the West Indians are good at a Sorting Test of Concept Formation and on some items from a battery of Piaget tests, very weak on a non-verbal Matrices and an Abstraction test, as well as on Kohs Blocks and other Piaget items—mainly those involving visualization and conservation.

It follows that the prospects for a culture-fair test or tests which might be used, for example, to test the educational potential of West Indian immigrant children in this country, are not good. Tests which yielded the best measures of *g* according to a factor analysis of West Indian results, were those given individually with the simplest instructions; and these tended to show as high correlations with environmental variables as any. In the writer's view, the most predictive instrument for immigrants is undoubtedly the WISC Verbal, provided the children are old enough and that it is given by a tester with some familiarity with West Indian speech. WISC Performance, Progressive Matrices and the non-verbal items of Terman-Merrill are less suitable in so far as success at these seems to depend on all sorts of complex factors such as insecure, female-dominated upbringing, and lack of play and spatial experience. Of the performance tests used in this research, the most suitable seemed to be Porteus Mazes and Draw-a-Man.

VII.—REFERENCES.

- BERNSTEIN, B. (1961). Social structure, language and learning. *Educ. Res.*, 3, 163-176.
- DAWSON, J. L. (1963). *Psychological Effects of Social Change in a West African Community*. Ph.D. Thesis, University of Edinburgh.
- GOLDSTEIN, H., and SCHEERER, M. (1941). Abstract and concrete behavior. *Psychol. Monogr.*, 53, No. 239, pp. 151.
- HUDSON, W. (1960). Pictorial depth perception in sub-cultural groups in Africa. *J. Soc. Psychol.*, 52, 183-208.
- IRVINE, S. H. (1964). *Testing Abilities and Attainments in Africa*. Paper read at Brit. Psychol. Soc. Educ. Sec. Meeting.
- JAHODA, G. (1956). Assessment of abstract behavior in a non-Western culture. *J. Abn. Soc. Psychol.*, 53, 237-243.
- MACARTHUR, R. S., IRVINE, S. H., and BRIMBLE, A. R. (1964). *The Northern Rhodesian Mental Ability Survey, 1963*. Lusaka: Rhodes-Livingstone Institute.
- MCCONNELL, J. (1954). Abstract behavior among the Tepehuan. *J. Abn. Soc. Psychol.*, 49, 109-110.
- MISCHEL, W. (1958). Father-absence and delay of gratification: cross-cultural comparisons. *J. Abn. Soc. Psychol.*, 56, 57-61.
- SCHWARZ, P. A. (1961). *Aptitude Tests for Use in the Developing Nations*. Pittsburgh, Pa.: American Institute for Research.
- SHAPIRO, M. B. (1954). An experimental investigation of the Block Design Rotation Effect. *Brit. J. Med. Psychol.*, 27, 84-88.
- VERNON, P. E. (1961). *Selection for Secondary Education in Jamaica*. Kingston, Jamaica: Government Printer.
- VERNON, P. E. (1962). Intellectual development in non-technological societies. *Proc. XIV. Internat. Congr. Appl. Psychol.*, II, 94-105.
- WITKIN, H. A. et al (1962). *Psychological Differentiation: Studies of Development*. New York: Wiley.

(Manuscript received 20th October, 1964.)

THE PSYCHOLOGY OF BEGINNING READING: AN EXPERIMENT WITH TWO METHODS

BY EMPERATRIZ S. TENSUAN AND FREDERICK B. DAVIS*

(Pasay City (Philippines) Division of City Schools and University of Pittsburgh)

SUMMARY. In one of the largest and best-controlled experiments in the teaching of beginning reading, data were obtained that lead to the conclusion that the combination, or multiple-approach method of teaching reading, yields better results in teaching Pilipino in grades 1 and 2 than the cartilla, or direct-phonetic-drill, method. Furthermore, the data strongly suggest that training in learning to read Pilipino by the combination method facilitates learning to read English more than learning to read Pilipino by the cartilla method.

Because English is less regular in its grapheme-phoneme correspondence than Pilipino, the cartilla method, or any method similar to it, is even less likely to be the best way to teach beginning reading in the United States. The data lend no support to arguments in favor of using direct phonic drill in teaching beginning reading.

I.—INTRODUCTION.

SEVERAL methods of introducing phonics in beginning reading instruction have been devised. Although many investigators have evaluated these methods experimentally, there is still substantial disagreement about how and when phonics should be employed. The question is more important in English-speaking countries than in many others because the correspondence between graphemes and phonemes is not close in English. In fact, the irregularities of English spelling have led to the development of alphabets like Unifon and Pitman's augmented Roman alphabet in which each grapheme has only one sound.

Even if these special orthographies are used, the problem of how to introduce phonics in beginning reading remains. The method recommended by most educational psychologists for teaching beginning reading makes use of multiple approaches and is referred to in this paper as the combination method. It stresses first the building of an apperceptive background so that pupils will know the meaning of all words that will occur in their early reading activities. Pictures, discussions, stories, and field trips play their part in developing this background. Second, it encourages the recognition of words and phrases as meaningful wholes. A pupil can associate an entire word or phrase with its pronunciation and its real-life meaning about as easily as he can associate a grapheme with a phoneme. Third, it systematically associates graphemes with their phonemes in words and phrases that the pupil is already reading, thereby keeping the sounding out and blending of phonemes in the normal flow of speech. From the point of view of those who advocate the combination method, beginning reading consists essentially of the following processes:

- (1) Associating the visual forms of familiar words and phrases with their real-life meanings.
- (2) Recognizing these familiar words and phrases in the material read.

* The senior author planned and carried out the study; the junior author analysed the data and wrote this paper. The authors gratefully acknowledge the assistance of the Philippine Center for Language Study, the Department of English of the University of California at Los Angeles, and the Institute of Language Arts, Teachers' College, Columbia University.

- (3) Identifying graphemes in familiar words and phrases and learning their sounds.
- (4) Identifying graphemes in unfamiliar words and phrases and pronouncing them (vocally or subvocally) in sequence.
- (5) Identifying these unfamiliar words and phrases.
- (6) Recognizing their real-life meanings.
- (7) Weaving together the meanings of both familiar and unfamiliar words and phrases to recreate the ideas of the writer.

A method widely used in the past and still advocated by a few educators puts great emphasis on beginning instruction in reading by teaching the child to associate graphemes and phonemes so that the sight of each grapheme evokes in him the corresponding phoneme, which he utters (or at least subvocalizes). A large chart with a list of the graphemes to be memorized used to be standard equipment in classrooms where reading was taught in this way. From the point of view of those who advocate this method, beginning reading consists essentially of the following processes :

- (1) Learning grapheme-phoneme associations.
- (2) Recognizing the graphemes to be read.
- (3) Pronouncing (vocally or subvocally) the corresponding phonemes in sequence.
- (4) Identifying the words sounded out.
- (5) Recognizing the real-life meanings of the words.
- (6) Weaving together these meanings to recreate the ideas of the writer.

This method can be expected to be most effective if the language being taught has a close correspondence of graphemes with phonemes. A language that meets this requirement is Pilipino, the national language of the Philippines, which is based on Tagalog, the predominant spoken language of the central part of the island of Luzon, including the city of Manila. For the purposes of this paper, we may regard Tagalog and Pilipino as interchangeable. If one knows the sounds given to letters (such as *a*, pronounced *ā* in Pilipino) and diphthongs (such as *ay*, pronounced *i*), he can sound out any word in Pilipino. As might be expected, it has long been supposed that the most efficient method of teaching beginning reading in Pilipino (or Tagalog) is the one just described, which in the Philippines is called the cartilla method.

To determine whether the cartilla or the combination method is the more effective for teaching Pilipino in grades 1 and 2, the Pasay City Schools, with the encouragement of the Bureau of Public Schools of the Philippine government, planned and carried out an experiment under the direction of the senior author. In the Philippines, children learn to read their native tongue in grade 1. In Manila and its vicinity, this is Pilipino. Simultaneously, the children are taught English as a second language. This is generally taught by some variant of the combination method. Beginning with grade 3, English becomes the language of the classroom ; subject matter is taught in English and English textbooks are used.

II.—THE SAMPLE.

Ten elementary schools in different sections of Pasay City participated in the experiment. Each school included at least one class taught by the cartilla method and one class taught by the combination method ; in seven schools, there

were more classes taught by the cartilla method than by the combination method. In all, there were twenty-one classes taught by the cartilla method and sixteen taught by the combination method. At the beginning of grade 1, there were 1,850 pupils in these classes: 1,050 taught by the cartilla method and 800 taught by the combination method. At the end of grade 1, the children who were promoted were kept in the same class groups in grades 2 and 3. The number of pupils for whom complete data were available was 1,149 at the end of grade 2. Of these, 658 were in classes taught Pilipino by the cartilla method and 491 in the classes taught by the combination method. At the end of grade 3, complete data were available for 952 pupils, of whom 528 were taught Pilipino by the cartilla method and 424 by the combination method. Drop-outs occurred as a result of absence at the time one or more of the tests were administered, of family moves, and of a wide variety of other factors.

III.—THE TEACHERS.

Each pupil who went through grades 1-3 had three teachers. All the teachers were known in advance to be competent; nonetheless, a composite rating based on length of experience, type of professional training, efficiency as judged by supervisors, and civil-service-eligibility score was computed for each one. These composite ratings were then summed for the two teachers each pupil had had at the end of grade 2 and for the three teachers he had had at the end of grade 3. The sums in each case were expressed as a single-digit Teacher Background Index in seven categories. The appropriate Teacher Background Index was assigned to each pupil and used as a covariate in analysis of covariance. Other data pertaining to the teachers in grades 1, 2 and 3 indicate that there were no significant differences with respect to salary or total income between those using the cartilla method and those using the combination method.

In grades 1 and 2, all teachers were provided with instructional materials and outlines of the procedures to be followed in teaching reading in Pilipino. This was done to maintain as much uniformity as possible in teaching methods within each of the two experimental groups. The teachers knew the purpose of the experiment and their prevailing opinion was that the cartilla method would prove the more effective way to teach reading in Pilipino.

IV.—THE MEASURES.

At the beginning of grade 1, each pupil was given the *Philippine Mental Ability Test*, Form II. It comprises fifty items categorized as follows:

Category	No. of Items
Information	5
Comparison	5
Classification	9
Perception	5
Omission	10
Absurdities	10
Logical Order	6

A raw score, consisting of the number of items marked correctly, was obtained for each pupil. Reading is not required in taking the test, which was administered in Pilipino.

At the end of grade 1, the *Philippine Achievement Test*, a battery of group tests given in Pilipino, was administered to all pupils. The battery consisted of a sixty-item arithmetic section, a thirty-five-item social-studies section, and a sixty-item reading section. Scores from this testing are not included in the analyses presented and are mentioned simply to indicate that it familiarized all pupils with the kinds of tests to be given a year later.

At the end of grade 2, an appropriate and different form of the *Philippine Achievement Test* was administered, consisting of a fifty-item social-studies section, a sixty-item arithmetic-problems section, a forty-item reading section, and an eighty-five item language-usage and spelling section. These are listed in Table 1. Raw scores on each of these sections, expressed as the number of

TABLE 1
IDENTIFICATION OF VARIABLES 21-28 IN TABLES 2 THROUGH 7.

Variable Number	Identification
21	Age in Grade 2, Expressed in Index Numbers.
22	Days in School in Grades 1 and 2.
23	Raw Score on Philippine Test of Mental Ability, Form II, Pilipino Edition.
24	Teacher Background Index (Experience, Training, Efficiency Rating, Civil-Service Eligibility) Expressed in Seven-Point Scale.
25	Raw Score on Philippine Achievement Test, Social Studies Section, Pilipino Edition.
26	Raw Score on Philippine Achievement Test, Arithmetic Problems Section, Pilipino Edition.
27	Raw Score on Philippine Achievement Test, Reading Comprehension Section, Pilipino Edition.
28	Raw Score on Philippine Achievement Test, Language Section, Pilipino Edition.

items marked correctly, have been used as separate criteria for judging the performance of the two groups of pupils—those taught by the cartilla method and those taught by the combination method. The tests were administered in Pilipino. In some respects, the reading-section scores constitute the most important criterion for judging the effectiveness of the two methods; however, a child's reading ability also determines to a considerable extent his scores on the other parts. In fact, it may be argued that skill in reading content in subject-matter fields is one of the principal objectives in the teaching of reading.

At the end of grade 3, an English edition of the *Philippine Achievement Test* was administered to all the pupils in the study. This covered the same subject-matter fields as the Pilipino edition administered at the end of grade 2.

V.—CHRONOLOGICAL AGE AND SCHOOL ATTENDANCE.

In the Philippine Islands, children normally enter grade 1 at age 7. To allow for the influence of age on performance, in reading, each pupil's chronological age was recorded at the end of grade 1. These ages, expressed as months, were coded by quarter years for use as a covariate in analysis of covariance. Similarly, the number of days of school attendance for each pupil in grades 1, 2 and 3 was obtained. In the analysis of the data obtained at the end of grade 2,

the number of days of school attendance for each pupil in grades 1 and 2 was used as a covariate. In the analysis of the data obtained at the end of grade 3, the number of days of school attendance in grades 1, 2 and 3 was used as a covariate.

VI.—TEACHING MATERIALS AND PROCEDURES ; SCHOOL FACILITIES AND EQUIPMENT.

The books, courses of study, guides, and other instructional materials used by the group taught Pilipino by the cartilla method and those used by the group taught by the combination method were the same except for the cartilla exercises (sounding out letters and diphthongs) and the reading-readiness exercises used by the combination method group. The cartilla exercises were taken from *Unang Hakbang sa Pagbasa (First Steps in Reading)* and were the same in all classes taught by that method. The reading-readiness exercises for the combination-method classes were taken from *Halinang Maglaro (Come, Let's Play)*. Both groups used the *Halina* series as readers.

The courses of study for all subject matter taught in grades 1, 2 and 3 were the same in the two groups. Time schedules accompanied these courses of study to insure that all teachers of both groups not only covered the same subject matter, but spent approximately equal amounts of time on its various subdivisions. Curriculum materials were prepared in advance of each year in workshops held for the purpose. It was judged that the school facilities and equipment available in each of the ten co-operating schools were very similar.

VII.—ANALYSIS OF THE DATA.

Table 1 lists the variables used in analyzing the scores obtained on the achievement test given at the end of grade 2. Tables, 2, 3 and 4 show the means, standard deviations, and intercorrelations of the eight variables in the combination group, the cartilla group, and the entire sample respectively

TABLE 2

GRADE 2 : MEANS, STANDARD DEVIATIONS, AND INTERCORRELATIONS OF VARIABLES 21-28
IN THE COMBINATION-METHOD GROUP (N=491).

Variable	21 ^a	22 ^b	23 ^c	24 ^d	25 ^c	26 ^c	27 ^c	28 ^c
21. Age		-.08	-.07	-.03	-.03	-.08	-.04	-.07
22. School attendance ..			-.09	-.09	-.12	-.05	-.08	-.08
23. Mental ability				-.05	-.36	-.34	-.33	-.30
24. Teacher index					-.23	-.17	-.11	-.11
25. Social studies						-.58	-.63	-.60
26. Arithmetic							-.57	-.62
27. Reading								-.65
28. Language								
Mean	8.81	371.36	31.94	4.57	43.18	37.80	35.23	43.48
S.D.	3.15	8.10	4.80	1.14	9.16	14.66	9.37	20.58

^a Expressed in index numbers (3 months per index number). 8.81 equals 9.29 years.

^b Expressed in days.

^c Expressed in raw scores.

^d Expressed in 7-point scale.

TABLE 3

GRADE 2: MEANS, STANDARD DEVIATIONS, AND INTERCORRELATIONS OF VARIABLES 21-28
IN THE CARTILLA-METHOD GROUP (N=658).

Variable	21 ^a	22 ^b	23 ^c	24 ^d	25 ^e	26 ^e	27 ^e	28 ^e
21. Age		-.10	.15	-.18	.09	.14	.06	.07
22. School attendance ..			.07	.25	.09	.03	.05	.09
23. Mental ability12	.37	.30	.39	.33
24. Teacher index17	.11	.14	.08
25. Social studies55	.66	.64
26. Arithmetic59	.61
27. Reading68
28. Language.....								
Mean	9.17	370.90	31.27	4.27	42.38	35.29	33.52	41.47
S.D.	3.70	8.64	5.38	1.50	9.55	14.39	9.85	19.76

TABLE 4

GRADE 2: MEANS, STANDARD DEVIATIONS, AND INTERCORRELATIONS OF VARIABLES 21-28
IN ENTIRE SAMPLE (N=1,149).

Variables	21 ^a	22 ^b	23 ^c	24 ^d	25 ^e	26 ^e	27 ^e	28 ^e
21. Age		-.09	.12	-.07	.06	.11	.05	.07
22. School attendance ..			.08	.13	.10	.04	.06	.09
23. Mental ability11	.37	.32	.37	.32
24. Teacher index19	.14	.14	.09
25. Social studies56	.65	.62
26. Arithmetic59	.61
27. Reading67
28. Language.....								
Mean	9.01	371.09	31.56	4.40	42.72	36.36	34.25	42.33
S.D.	3.48	8.41	5.15	1.37	9.39	14.56	9.68	20.13

^a Expressed in index numbers (3 months per index number). 9.17 equals 9.38 years.
9.01 equals 9.34 years.

^b Expressed in days.

^c Expressed in raw scores.

^d Expressed in 7-point scale.

The average chronological age was slightly lower in the combination group than in the cartilla group, but the difference is not significant at the .05 level. The average number of days of school attendance was almost identical in the two groups, the combination group having slightly, but insignificantly, the better record.

At the beginning of grade 1, the combination group was slightly superior with respect to mental-test score. The difference of .67 raw-score point is significant at the .05 level. The average Teacher Background Index of the first- and second-grade teachers in the combination group was higher than that in the cartilla group, the difference being significant at the .001 level. Of the four variables (age, days in school attendance, mental-test score, and teacher background) the influence of which it was desired to eliminate by adjusting the

mean scores of the two groups in each of variables 25 through 28 (social studies, arithmetic, reading, and language), the combination group had the higher scores in three.

The patterns of intercorrelations in Tables 2 and 3 are similar; the most striking discrepancy is that between $-.09$ and $.25$, expressing the relationship between variables 22 (days in school attendance) and 24 (teacher background) in the combination and cartilla groups, respectively. This difference is statistically significant but no plausible hypothesis to explain it suggests itself. In general, the intercorrelations of the adjusting variables (21-24) are close to zero. The correlations of variables 21 and 22 with the criterion variables (25-28) are also low. Variable 24 shows slightly higher relationships to the criterion variables. The mental-test scores (variable 23) have correlations between $.30$ and $.40$ with the four criterion variables. It had been expected that these relationships would be somewhat higher. Apparently, the *Philippine Mental Ability Test*, given at the beginning of grade 1, is not strongly predictive of academic achievement at the end of grade 2.

Multiple correlations of variables 21-24 with each of the criterion variables are shown in Table 5, together with the corresponding regression coefficients

TABLE 5

GRADE 2: REGRESSION COEFFICIENTS IN RAW-SCORE AND STANDARD-SCORE FORM; AND MULTIPLE CORRELATIONS OF ADJUSTING VARIABLES 21-24 WITH VARIABLES 25, 26, 27 AND 28.

	Multiple Correlation	Regression Coefficient	
		Raw-Score Form	Standard-Score Form
$R_{25 \cdot 21, 22, 23, 24} \dots$.40	.11	.04
$b_{25, 21 \cdot 22, 23, 24} \dots$.07 ^a	.06 ^a
$b_{25, 22 \cdot 21, 23, 24} \dots$.62 ^c	.34 ^c
$b_{25, 23 \cdot 21, 22, 24} \dots$		1.02 ^c	.15 ^c
$b_{25, 24 \cdot 21, 22, 23} \dots$			
$R_{26 \cdot 21, 22, 23, 24} \dots$.35	.35 ^b	.08 ^b
$b_{26, 21 \cdot 22, 23, 24} \dots$.01	.01
$b_{26, 22 \cdot 21, 23, 24} \dots$.85 ^c	.30 ^c
$b_{26, 23 \cdot 21, 22, 24} \dots$		1.17 ^c	.11 ^c
$b_{26, 24 \cdot 21, 22, 23} \dots$			
$R_{27 \cdot 21, 22, 23, 24} \dots$.38	.04	.01
$b_{27, 21 \cdot 22, 23, 24} \dots$.03	.03
$b_{27, 22 \cdot 21, 23, 24} \dots$.66 ^c	.35 ^c
$b_{27, 23 \cdot 21, 22, 24} \dots$.67 ^c	.09 ^c
$b_{27, 24 \cdot 21, 22, 23} \dots$			
$R_{28 \cdot 21, 22, 23, 24} \dots$.33	.24	.04
$b_{28, 21 \cdot 22, 23, 24} \dots$.14 ^a	.06 ^a
$b_{28, 22 \cdot 21, 23, 24} \dots$		1.20 ^c	.31 ^c
$b_{28, 23 \cdot 21, 22, 24} \dots$.84 ^a	.06 ^a
$b_{28, 24 \cdot 21, 22, 23} \dots$			

^a Significantly different from zero at the .05 level.

^b Significantly different from zero at the .01 level.

^c Significantly different from zero at the .001 level.

in both raw-score and standard-score form. In each case, the multiple correlation of the adjusting variables (21-24) with the criterion variable is positive, significant, and large enough to warrant practical use in analysis of covariance. The results of such an analysis are shown in Tables 6 and 7. For each of the

TABLE 6

GRADE 2: RAW-SCORE MEANS AND MEANS AFTER ADJUSTMENT BY VARIABLES 21-24.

Variable	Scores	Combination- Method Group	Cartilla-Method Group	Mean Difference
25	Raw	43.18	42.38	.80
	Adjusted	42.77	42.68	.09
26	Raw	37.80	35.29	2.51
	Adjusted	37.36	35.61	1.75
27	Raw	35.23	33.52	1.71
	Adjusted	34.87	33.78	1.09
28	Raw	43.48	41.47	2.01
	Adjusted	42.90	41.90	1.00

TABLE 7

GRADE 2: ANALYSIS OF COVARIANCE OF PERFORMANCE ON FOUR CRITERION VARIABLES OF PUPILS IN COMBINATION-METHOD AND CARTILLA-METHOD GROUPS.

Variable	Source of Variation	Sum of Squares for Error	dof	Mean Square	F
25	Total	84,904.27	1,444		
	Within	84,902.05	1,143		
	Adjusted Means....	2.22	1	74.28 2.22	.03
26	Total	213,806.46	1,144		
	Within	212,959.19	1,143		
	Adjusted means....	847.27	1	186.32 847.27	4.55
27	Total	92,142.24	1,144		
	Within	91,814.02	1,143		
	Adjusted Means....	328.22	1	80.33 328.22	4.09
28	Total	413,303.13	1,144		
	Within	413,028.95	1,143		
	Adjusted Means....	274.18	1	361.36 274.18	.76

four criterion variables, the difference between adjusted-score means is smaller than that between raw-score means. In each case, the combination-method group obtained the higher adjusted means core. The adjusted means constitute our best estimates of the average achievement of the pupils when the influences of variables 21-24 on their achievement have been allowed for. Of the four differences between adjusted means, only those for arithmetic and reading

are significant at the .05 level. Thus, we may conclude that, under the circumstances described, it is highly likely that pupils learn to read Pilipino slightly more effectively in grades 1 and 2 when taught by the combination method than when taught by the cartilla method. Since Pilipino is a language in which the cartilla method should be at its best, it is reasonable to infer that if it is not superior in teaching Pilipino it can hardly be expected to be superior in teaching English.

Table 8 shows the variables involved in comparing the performance in *English* of pupils at the end of grade 3 who were taught to read Pilipino by the combination and cartilla methods, respectively. Data analogous to those in Tables 2-4 are shown in Tables 9, 10, and 11. In general, the correlation

TABLE 8
IDENTIFICATION OF VARIABLES 31-38 IN TABLES 9 THROUGH 14.

Variable Number	Identification
31	Age in Grade 2, Expressed in Index Numbers.
32	Days in School in Grades 1, 2, and 3.
33	Raw Score on Philippine Test of Mental Ability, Form II, Pilipino Edition.
34	Teacher Background Index (Experience, Training, Efficiency Rating, Civil-Service Eligibility) Expressed in Seven-Point Scale.
35	Raw Score on Philippine Achievement Test, Social Studies Section, English Edition.
36	Raw Score on Philippine Achievement Test, Arithmetic Problems Section, English Edition.
37	Raw Score on Philippine Achievement Test, Reading Comprehension Section, English Edition.
38	Raw Score on Philippine Achievement Test, Language Section English Edition.

TABLE 9
GRADE 3: MEANS, STANDARD DEVIATIONS, AND INTERCORRELATIONS OF VARIABLES 31-38 IN THE COMBINATION-METHOD GROUP (N=424).

Variable	31 ^a	32 ^b	33 ^c	34 ^d	35 ^c	36 ^c	37 ^c	38 ^c
31. Age		-.05	.12	-.00	.11	-.07	-.02	.01
32. School attendance ..			.10	.02	.24	.06	.22	.18
33. Mental ability06	.25	.30	.29	.21
34. Teacher index					-.04	-.02	.09	.14
35. Social studies52	.63	.59
36. Arithmetic47	.51
37. Reading61
38. Language								
Mean	8.74	552.15	31.97	4.19	37.64	31.17	54.85	28.24
S.D.	3.10	11.72	4.77	1.36	15.93	11.96	13.69	15.42

^a Expressed in index numbers (3 months per index number). 8.74 equals 9.27 years.

^b Expressed in days.

^c Expressed in raw scores.

^d Expressed in 7-point scale.

TABLE 10

GRADE 3: MEANS, STANDARD DEVIATIONS, AND INTERCORRELATIONS OF VARIABLES 31-38
IN THE CARTILLA-METHOD GROUP (N=528).

Variables	31 ^a	32 ^b	33 ^c	34 ^d	35 ^e	36 ^e	37	38 ^e
31. Age		-.19	.15	-.10	-.04	-.08	-.02	-.05
32. School attendance ..			.10	.30	.22	-.04	.16	.16
33. Mental ability16	.33	.31	.29	.30
34. Teacher index22	.14	.25	.18
35. Social studies50	.68	.62
36. Arithmetic52	.52
37. Reading56
38. Language.....								
Mean	9.04	551.41	31.41	3.70	38.09	30.24	52.07	25.82
S.D.	3.23	12.55	5.41	1.57	15.23	11.47	14.60	14.68

TABLE 11

GRADE 3: MEANS, STANDARD DEVIATIONS, AND INTERCORRELATIONS OF VARIABLES 31-38
IN ENTIRE SAMPLE (N=952).

Variable	31 ^a	32 ^b	33 ^c	34 ^d	35 ^e	36 ^e	37 ^e	38 ^e
31. Age		-.13	.14	-.97	-.07	-.07	-.00	-.03
32. School attendance ..			.10	.19	.23	.05	.19	.17
33. Mental ability13	.29	.31	.29	.26
34. Teacher index10	.07	.20	.17
35. Social studies51	.65	.60
36. Arithmetic50	.52
37. Reading58
38. Language.....								
Mean	8.91	551.74	31.66	3.91	37.89	30.66	53.31	26.90
S.D.	3.18	12.18	5.14	1.50	15.54	11.69	14.26	15.06

^a Expressed in index numbers (3 months per index number). 9.04 equals 9.35 years.
8.91 equals 9.32 years.

^b Expressed in days.

^c Expressed in raw scores.

^d Expressed in 7-point scale.

patterns are similar. As would be expected, the *Philippine Test of Mental Ability*, administered in Pilipino in grade 1, has slightly lower validity for predicting achievement in English at the end of grade 3 than achievement in Pilipino at the end of grade 2.

The multiple correlations of variables 31-34 with three of the four criteria (variables 35-38) are lower than those found in grade 2. These are shown in Table 12.

The raw-score and adjusted means for pupils in the combination-method and cartilla-method groups on each of the four criterion variables are presented in Table 13 and the significance of the differences among the adjusted means are given in Table 14. With respect to variable 35 (social studies), the cartilla group was slightly and insignificantly higher. On the other three variables, the

TABLE 12

GRADE 3: REGRESSION COEFFICIENTS IN RAW-SCORE AND STANDARD-SCORE FORM; AND MULTIPLE CORRELATIONS OF ADJUSTING VARIABLES 31-34 WITH VARIABLES 35, 36, 37 and 38.

	Multiple Correlation	Regression Coefficient	
		Raw-Score Form	Standard-Score Form
$r_{35 \cdot 31, 32, 33, 34} \dots$.36		
$b_{35, 31 \cdot 32, 33, 34} \dots$.32 ^a	.07 ^a
$b_{35, 32 \cdot 31, 33, 34} \dots$.26 ^c	.20 ^c
$b_{35, 33 \cdot 31, 32, 34} \dots$.79 ^a	.26 ^c
$b_{35, 34 \cdot 31, 32, 33} \dots$.38	.04
$r_{36 \cdot 31, 32, 33, 34} \dots$.31		
$b_{36, 31 \cdot 32, 33, 34} \dots$.12	.03
$b_{36, 32 \cdot 31, 33, 34} \dots$.02	.02
$b_{36, 33 \cdot 31, 32, 34} \dots$.68 ^c	.30 ^c
$b_{36, 34 \cdot 31, 32, 33} \dots$.28	.04
$r_{37 \cdot 31, 32, 33, 34} \dots$.36		
$b_{37, 31 \cdot 32, 33, 34} \dots$		-.04	-.01
$b_{37, 32 \cdot 31, 33, 34} \dots$.16 ^c	.14 ^c
$b_{37, 33 \cdot 31, 32, 34} \dots$.72 ^a	.26 ^c
$b_{37, 34 \cdot 31, 32, 33} \dots$.133 ^c	.14 ^c
$r_{38 \cdot 31, 32, 33, 34} \dots$.33		
$b_{38, 31 \cdot 32, 33, 34} \dots$		-.18	-.04
$b_{38, 32 \cdot 31, 33, 34} \dots$.15 ^c	.12 ^c
$b_{38, 33 \cdot 31, 32, 34} \dots$.71 ^c	.24 ^c
$b_{38, 34 \cdot 31, 32, 33} \dots$		1.18 ^c	.12 ^c

a Significantly different from zero at the .05 level.

b Significantly different from zero at the .01 level.

c Significantly different from zero at the .001 level.

TABLE 13

GRADE 3: RAW-SCORE MEANS AND MEANS AFTER ADJUSTMENT BY VARIABLES 31-34.

Variable	Scores	Combination-Method Group	Cartilla-Method Group	Mean Group Difference
35	Raw	37.64	38.09	-.45
	Adjusted	37.22	38.43	-1.21
36	Raw	31.17	30.24	.93
	Adjusted	30.91	30.45	.46
37	Raw	54.85	52.07	2.78
	Adjusted	54.22	52.58	1.64
38	Raw	28.24	25.82	2.42
	Adjusted	27.63	26.31	1.32

TABLE 14

GRADE 3 : ANALYSIS OF COVARIANCE OF PERFORMANCE ON FOUR CRITERION VARIABLES OF PUPILS IN COMBINATION-METHOD AND CARTILLA-METHOD GROUPS.

Variable	Source of Variation	Sum of Squares for Error	dof	Mean Square	F
35	Total	199,391.56	947		
	Within	199,061.47	946	210.42	
	Adjusted Means....	330.09	1	330.09	1.57
36	Total	117,221.23	947		
	Within	117,174.47	946	123.86	
	Adjusted Means....	46.76	1	46.76	.38
37	Total	168,529.99	947		
	Within	167,914.05	946	177.50	
	Adjusted Means ..	615.94	1	615.94	3.47
38	Total	192,651.16	947		
	Within	192,254.38	946	203.23	
	Adjusted Means....	396.78	1	396.78	1.95

combination group obtained insignificantly higher scores than the cartilla group. Of these, only the difference between the two groups in reading approaches significance at the .05 level. These data suggest that the type of training involved in teaching beginning reading in Pilipino by the combination method facilities learning to read English more than the type of training involved in teaching beginning reading in Pilipino by the cartilla method.

VIII.—DISCUSSION OF FINDINGS.

Under the conditions of this study, we must conclude that the practical difference resulting from teaching Philippine children to read Pilipino by the combination or the cartilla method is small. What difference there is favours the combination method. Since Pilipino is a language in which there is close correspondence of graphemes and phonemes, the cartilla method should have its maximum value in that language. In other languages, such as English, in which the correspondence of graphemes and phonemes is less, the cartilla method, or any method like it, is at a disadvantage. Children find that they cannot pronounce a large number of words correctly by using the grapheme-phoneme associations that they have been taught ; as a result, they sometimes lose confidence in the techniques of word attack that they have learned, in the teachers who taught them these techniques, and in their own ability to learn to read.

The implication is plain that the combination method will be found to yield better results in teaching beginning reading in English than the cartilla method, or any method similar to it. It must be pointed out that this implication does not suggest that phonics should not be taught in beginning reading. The combination method does include the teaching of phonics and of sound-blending techniques in a functional way after an initial sight vocabulary has been developed. It is probable that many variations of the teaching of phonics in the combination method would prove to be about equally effective. The data of this study lend no support to the arguments put forth by those who advocate direct phonic drill on the sounds of letters and diphthongs during the first few

weeks of reading instruction. This should not be taken to mean, however, that facility in associating graphemes and phonemes is of little consequence in beginning reading.

It should be emphasized that the criteria by which performance in reading was judged at the end of grade 2 and of grade 3 were tests of paragraph comprehension in silent reading and tests of subject-matter achievement. If the criteria had been tests of ability to pronounce isolated words or nonsense syllables at first sight, it is quite possible that the combination method would not have been found superior. It is commonly observed that first-grade pupils taught by the cartilla method very quickly become able to pronounce correctly most words in Pilipino that are presented to them, including words they have never seen before and whose meanings they do not know. This rapid acquisition of the ability to pronounce words at first sight is more impressive to laymen, and apparently to some linguists and teachers, than it deserves to be. It is not always evidence of a corresponding ability to comprehend. And it is the latter that constitutes the essential core of the agglomeration of skills that we call reading. This should not be taken to mean, however, that facility in associating graphemes and phonemes is of little consequence in the teaching of beginning reading.

(Manuscript received June, 1964.)

PERSONALITY FACTORS IN TEACHER TRAINEE SELECTION

By SISTER M. SIMEON TARPEY

(*Department of Psychology, University College, Dublin*)

SUMMARY. A personality questionnaire, intelligence test, teacher attitude inventory and interest inventory were administered to 128 students in four different training colleges, with the following results :

(1) There was no significant difference in the intelligence level of the four colleges and no significant correlation between intelligence test scores and the teaching mark.

(2) Though the general personality pattern was normal, significant correlation between personality factors and the teaching mark was found in only one college.

(3) Scores on the teacher attitude inventory were lower than the American norms. Positive correlation, however, was found between the inventory and the teaching mark of all four colleges, a correlation which was significant for two. Unexpected correlations were also found between the inventory and some personality factors.

(4) Scores were high on outdoor-art-music-literary-social service interests, but few significant correlations with teaching marks were found.

(5) As the personality and interest profiles appeared to be normal and expected, it is suggested that the teaching mark, the basic criterion with which they were correlated, is not in itself a reliable *statistical* measure.

I.—INTRODUCTION.

BIBLIOGRAPHIES showing the research done in the field of teacher training over the past fifty years list well over 1,000 titles (Domas and Tiedeman, 1950 ; Evans, 1961). Most of this research, however, has been carried out in Great Britain and in America, and though the results of British and American research are not always applicable in Ireland, where educational conditions are somewhat different, it was thought opportune to try to investigate the Irish situation in the light of these findings. In Ireland, in the smaller towns and villages and rural areas especially, teachers generally hold positions of honour and authority and are a potent influence in the formation of the younger generations. Because of the importance of having trained teachers in the schools, entrance to training colleges is made comparatively easy by the provision of grants and scholarships, awarded in the main on the results of school certificate examinations. Recently, interviews have been introduced but little approach has been made towards selection on a scientific basis.

II.—PLAN OF INVESTIGATION.

(i) *Subjects and Materials.*

Three Irish training colleges were selected for the investigation, a fourth college, in England, being used for comparison :

College 1 : a three-year college for women, unaided, independent, fee-paying. Twenty-eight students, the full third-year group.

College 2 : a men's college, two-year training with Government grants. Thirty-one students, a random sample of the final year group.

College 3 : a women's college under the Ministry of Education of Northern Ireland, three-year course. Thirty-nine volunteers from the final year students.

College 4 : a mixed college, three-year training, in an English industrial city. Thirty students from the primary teachers' group.

This gave a total of 128 students : ninety-eight from the Irish colleges, sixty-seven women and thirty-one men ; and thirty from the English college, twenty-six women and four men.

Since research into the qualities and characteristics of a good teacher and the conditions which determine or are associated with efficient teaching (Cattell, 1931 ; Birkinshaw, 1935 ; Evans, 1952, 1953, 1957 ; Vernon, 1953, Warburton, *et al.*, 1963, Allen, 1963, and others) suggests the importance of the teacher's personality, general intellectual ability, attitudes to children and to teaching, and avocational interests, it was decided that the following standardised tests should be given to the students :

- (a) Cattell's Sixteen Personality Factor Questionnaire.
- (b) AH5 Group Intelligence Test.
- (c) Minnesota Teacher Attitude Inventory.
- (d) Kuder Preference Record—vocational.

A questionnaire : Motives for choice of teaching as a career, based on the research of Valentine (1934) and Tudhope (1944) and including as well some motives which it was thought might have particular relevance to teaching in Ireland, was also given (Tarpey, 1963) but is not being discussed in the present article.

(ii) *Experimental Procedure.*

These tests were administered between February and June, 1963. Because the English and North of Ireland students were already familiar with the testing situation, the AH4 was given to the Dublin students beforehand, as a practice test, otherwise the procedure was the same in all four colleges. The total results, with the addition of the final college teaching mark were then tabulated and a product-moment correlation programme was prepared for the electronic computer in the Physics Department of University College, Dublin. In order to enter the 16 P-F Questionnaire as a unit, it was necessary to drop one column of scores : Factor B was omitted as an intelligence score had already been obtained on the AH5. This entailed a certain loss, however, as Cattell points out in the Manual that the B score will not correlate fully with any intelligence test given under speeded conditions, the B factor being a measure of power rather than of speed. It was noted that the only student who obtained the full score on Factor B had a score of 19, E on the AH5 ; and that the one A student on the AH5 scored 10 on Factor B.

This gave a total of twenty-eight columns, the test results of 128 students. The variables included in the research were as follows :

- (1) Intelligence Test—AH5 : total score.
- (2) Minnesota Teacher Attitude Inventory—raw scores.
- (3) Teaching Mark : a numerical approximation to the literal mark.
- (4) Kuder Preference Record : Outdoor Interests (percentiles).
- (5) Kuder Preference Records : Mechanical Interests (percentiles).
- (6) Kuder Preference Record : Computational Interests (percentiles).
- (7) Kuder Preference Record : Scientific Interests (percentiles).
- (8) Kuder Preference Record : Persuasive Interests (percentiles).

- (9) Kuder Preference Record : Artistic Interests (percentiles).
- (10) Kuder Preference Record : Literary Interests (percentiles).
- (11) Kuder Preference Record : Musical Interests (percentiles).
- (12) Kuder Preference Record : Social Service Interests (percentiles).
- (13) Kuder Preference Record : Clerical Interests (percentiles).
- (14) Sixteen Personality Factors : A Cyclothymia *v.* Schizothymia (raw scores).
- (15) Sixteen Personality Factors : C Ego Strength (raw scores).
- (16) Sixteen Personality Factors : E Dominance (raw scores).
- (17) Sixteen Personality Factors : F Surgency (raw scores).
- (18) Sixteen Personality Factors : Conscientiousness (raw scores).
- (19) Sixteen Personality Factors : H Extraversion (raw scores).
- (20) Sixteen Personality Factors : I Sensitivity (raw scores).
- (21) Sixteen Personality Factors : L Paranoid Tendency (raw scores).
- (22) Sixteen Personality Factors : M Unconventionality (raw scores).
- (23) Sixteen Personality Factors : N Sophistication (raw scores).
- (24) Sixteen Personality Factors : O Anxiety (raw scores).
- (25) Sixteen Personality Factors : Q1 Radicalism (raw scores).
- (26) Sixteen Personality Factors : Q2 Self-sufficiency (raw scores).
- (27) Sixteen Personality Factors : Q3 Self-control (raw scores).
- (28) Sixteen Personality Factors : Q4 Id Pressure (tenseness) (raw scores).

The purpose of the investigation was to find out the degree of association between the different variables, therefore the mean scores and standard deviations were obtained, and product-moment correlations between Nos. 1, 2, 3, and 4-13, inclusive, and between Nos. 1, 2, 3, and 14-28, inclusive. Since the colleges did not belong to the same area, and were quite independent of each other, there was no guarantee that the teaching marks awarded in each college were of the same standard or meant the same thing; the results of each college were, therefore, kept separate. It was thought that because of the comparatively easy entry to the training colleges and because of the large number of grants and scholarships available, there might be socio-economic differences among the students. The total Irish group was, therefore, classified into Social Groups, according to the 1960 Classification of Occupations, and the results were compared with those of the English College. Table 1 gives a summary in percentages of the home-backgrounds of the students in the four colleges.

TABLE 1

OCCUPATIONAL CLASSIFICATION OF THE STUDENTS IN THE FOUR COLLEGES.

Social Group	College 1	College 2	College 3	College 4
I. Professional O	% 7.1	% —	% 5.1	% 10
II. Intermediate O.....	67.9	48.4	48.7	46.6
III. Skilled O.	21.4	32.2	30.8	26.6
IV. Partly-skilled O. *....	3.6	19.3	10.3	16.6
V. Unskilled O.	—	—	5.1	—

The samples for Groups I, IV and V were not satisfactory statistically, but they give some indication of trends. English groups III and IV were combined throughout. The mean scores, standard deviations and product-moment correlations were again obtained as for the colleges.

III.—STATEMENT OF RESULTS.

The findings of the present investigation corroborate the evidence of earlier research in other countries and emphasise its relevance to work in Ireland. No one pattern of successful teacher was found and there were appreciable differences among the students of the different training colleges.

(i) *Intelligence.*

Table 2 shows the mean AH5 score and the standard deviation of each of the four colleges and each of the social groups.

TABLE 2
MEAN SCORE AND STANDARD DEVIATION ON AH5 INTELLIGENCE TEST.

College	N	Mean	S.D.
1	28	26.25	7.174
2	31	32.741	8.381
3	39	31.051	7.448
4	30	31.900	6.957
Social Group	N	Mean	S.D.
II. (Irish)	53	29.622	8.812
III. (Irish)	30	30.933	7.469
IV. (Irish)	9	29.770	5.902
II. (English)	14	31.642	7.470
III and IV. (English)	13	32.384	6.604

Though the difference in the means of the different colleges and social groups was slight, the trend was not unexpected. Colleges 2 and 3 are the Irish grant-aided colleges, where school certificate results are the most heavily weighted factors in selection; with a good school certificate entrance is comparatively easy. In these samples at least the results of the school certificate examination would seem to be a fairly stable indicator of general intelligence. In College 1, where less weight is given to school certificate qualifications, the AH5 score was somewhat lower. Contrary to what was expected, no distinctive differences were found among the social groups.

The mean AH5 scores placed these students below the middle of a university student population, in the D and E groups, though the entire range of scores was 10–57, as against the university students' range of 13–62, as given in the AH5 Test Manual.

CORRELATION OF INTELLIGENCE WITH TEACHING MARK.

Correlations between intelligence test results and the teaching mark, both positive and negative, were found to be insignificant in all the colleges and in all the social groups. There was a difference, however, between the mean scores of the six A teachers and that of the four failures in teaching: the mean score of the former group was 28.3, that of the latter 24.5.

(ii) *Personality.*

The graphed results of the fifteen Personality Factors (Tarpey, 1963) gave a clear picture of the wide variation in personality among the students of the different colleges, again confirming the evidence of earlier research. In spite of

this great variety, however, the general pattern was normal and certain points of agreement were noted :

High Scores : (a) on Factor C, emotional stability. " Individuals having to adjust to difficulties thrown upon them from outside, e.g., teachers, score above average on C." (Manual). (b) On Factor F, surgency. High F scores indicate gay, enthusiastic personalities.

Low Scores : (a) On Factor L. L- suggests adaptable, trustful, tolerant people. (b) On Factor N. N- indicates naïve, simple, forthright individuals. (c) On Q1, indicating conservatism.

Average Scores : (a) On Factor A, cyclothymia *v.* schizothymia. (b) On G, superego strength, conscientiousness. (c) On O, an anxiety measure.

College 1 had the highest mean score on C, emotional stability and the lowest on E, suggesting submissiveness, and on Q1, suggesting conservatism, scores which correlated significantly and positively.

College 2, the men's college, was lowest on O, anxiety, and on Q4, in pressure, factors which correlated positively and significantly in all groups ; lowest also on I, indicating realism, and on M, indicating practicalness, but highest on E, dominance and on L-suggesting adaptability.

College 3 had the highest score on F, surgency, and on Q2, self-sufficiency, though this score was still on the lower end of the scale ; and was lowest on H-, suggesting shyness, L- suggesting adaptability and N- indicating naïvete.

College 4, the group of English students, seemed to be the least stable emotionally, but scored highest on I, indicating tendermindedness ; on O, anxiety ; on M indicating unconventionality ; on Q4 in pressure ; and on H adventurousness.

Social Groups : The scores for the English groups were on the whole much higher than those for the Irish groups. Irish Social Group IV and English Groups III and IV combined, both seemed to be less stable, more assertive and aggressive and yet more conservative than the others. English Group IV also showed as more bohemian, more casual and more expedient, differences in the Irish groups on the whole being less pronounced. The samples for Group IV, however, were too small for any valid generalisations to be made.

CORRELATION OF PERSONALITY FACTORS WITH THE TEACHING MARK.

In spite of the apparently good personality pattern, little correlation was found between personality factors and the teaching mark. In Colleges 2, 3 and 4, there was no significant correlation whatsoever of any personality factor with the teaching mark of these colleges. Table 3 shows the few significant correlations there were for College 1 and for the Social Groups.

TABLE 3

SIGNIFICANT CORRELATIONS OF PERSONALITY FACTORS WITH TEACHING MARK.

College 1 N=28	I. Social Group II N=53	I. Social Group III N=30
Factor A, $r = -.442$ Factor H-, $r = -.408$ Factor G, $r = +.446$ Factor M, $r = -.372$	Factor A, $r = -.302$ Factor I, $r = -.329$	Factor N, $r = +.360$ Factor Q3, $r = +.382$

Factor A measures the warm, sociable, easy-going person as contrasted with the more aloof, cold and rigid person, cyclothymia *versus* schizothymia. Cattell notes in the Manual that the highest ranking occupations on the A Factor are teaching and salesmanship; teachers who have to adapt cheerfully to a lot of compromises with human failings and to take a ceaseless impact of emotional problems score high. The *H Factor* measures something similar to the A Factor: the overt, adventurous, uninhibited person as contrasted with the shy, timid, restrained person. These two factors correlated, positively and significantly in College 1. The mean score for Factor A was 11.642, S.D. 2.767 and for Factor H, 9.5, S.D. 4.204, i.e., average A and H-.

It was not surprising to find a negative correlation between H- and the teaching mark: the H- individual reports himself to be intensely shy, convinced of his inferiority, slow and impeded in expressing himself, disliking occupations with personal contacts, preferring one or two close friends to large groups. It was surprising, however, to find a negative correlation between the A factor and the teaching mark. This may perhaps be explained, however, by the unusual 10 per cent. failure in this group, an occurrence which must necessarily distort the correlations. One of the failures scored 5 on the H factor, another 5 on G, but all had high scores on Factor A: 14, 15 and 17, while the average score of the A and B+ teachers was only 10.

Correlation of the teaching mark with *Factor G* was expected and was also found by Hadley; as reported in the Manual: "The G person regards himself as correct in and a guardian of manners and morals."

Factor M shows a subtle personality pattern which, as the Manual says, is rather difficult to define. "The M+ person has intense subjectivity and inner mental life. Although cheerful and irresponsible about practical matters, he actually has high internal spasmodic anxiety tensions." Though the mean score of College 1 on this factor is not high, the correlation shows that the M+ persons got the higher teaching marks. Cattell and Drevdagh (Manual) found this factor to distinguish the more creative artists and research workers. It also seems to distinguish the more creative teachers.

In Social Group II, two negative correlations were found: with Factor A, $r = -.302$, and with *Factor I-*, $r = -.329$. Factor I- represents a "sort of tough, masculine, practical, mature, realistic (no-nonsense) temperamental dimension." This correlation agrees with the findings of Warbuerton, *et al.* (1963), where I+, tendermindedness, sensitivity, was positively correlated with success in teaching: here there is negative correlation with I-.

In Social Group III, positive and significant correlations were found for *Factors N-* and Q3. Hadley (Manual) found that high N correlated negatively with teaching success. This is a dimension "associated with a generalised mental alertness, health and efficiency, a factor tending to be significantly low in both the major forms of psychosis and in neurosis." Even so, the Manual suggests that there is too much efficiency in N+ individuals to tolerate people and their failings, and perhaps more natural warmth and liking for people in N- individuals. A high Q3 score suggests a well-integrated balanced personality, showing self-control, persistence, foresight, considerateness of others, conscientiousness, qualities which are normally associated with teaching efficiency.

(iii) *Minnesota Teacher Attitude Inventory.*

The statistics here were worked on the raw scores, as it was felt that the percentiles were too closely grouped and involved a slight loss of information. The scores, ranging from -38 to 97, were found to be considerably lower than

the American norms, a result which agrees with the findings of Evans (1938). Table 4 shows the wide divergence in the scores of the different colleges and groups.

TABLE 4
MEAN SCORES AND STANDARD DEVIATIONS ON M.T.A.I.

College	N	Mean	S.D.
1	28	33.214	21.082
2	31	10.580	26.935
3	39	40.923	23.982
4	30	35.100	30.867
Social Group	N	Mean	S.D.
II. (Irish)	53	24.490	24.729
III. (Irish)	30	34.990	27.761
IV. (Irish)	9	14.220	19.210
II. (English)	14	34.428	25.858
III and IV. (English)	13	33.650	35.861

CORRELATION WITH THE TEACHING MARKS AND OTHER SCORES.

A most interesting result was the correlation of the Inventory with the teaching marks of colleges 2 and 3: *it was the only positive and significant correlation of the whole battery of tests with these two teaching marks.* College 2: $N=31$, $r=+.434$; college 3: $N=39$, $r=+.330$. Table 5 shows the other variables with which the Inventory correlated significantly.

TABLE 5
OTHER SIGNIFICANT CORRELATIONS WITH M.T.A.I.

College 2: N=31	Personality Factor A	$r=+.374$
	Personality Factor O	$r=+.366$
	Personality Factor Q4	$r=+.357$
	Personality Factor C	$r=-.406$
	Kuder Pref. Outdoor Interest	$r=+.365$
College 1: N=28	Kuder Pref. Literary Interest	$r=+.440$
	Personality Factor O	$r=-.539$
Social Group III: N=14	Personality Factor L	$r=-.434$
	Personality Factor Q4	$r=-.488$

No other significant correlations were found in college 3 and none at all in college 4. No significant correlations were found between Intelligence and Attitude Inventory scores.

These findings are difficult to interpret. The teaching mark of college 2 itself is not significantly correlated with any of the personality factors, and yet there is positive and significant correlation between the Attitude Inventory and the teaching mark, and between the Inventory and three personality factors.

The mean score for college 2 on Factor A is 11.29, S.D. 2.412; for O, 11.709, S.D. 3.611; and for Q4, 11.096, S.D. 4.913, the lowest scores of all four groups on these factors, O and Q4 showing internal correlation, $r = +.470$. Lower O and Q4 scores indicate placid, self-assured, confident, serene, unfrustrated individuals, while lower A scores indicate cool, reserved and detached people; the correlations here are not unexpected. The mean score for C is higher, however: 14.935, S.D. 3.775. This correlates negatively and significantly with O, $r = -.415$, and with Q4, $r = -.626$. The correlation with the Inventory is $r = -.406$.

Does this last correlation suggest that emotional stability is not an important measure in the teacher attitude scale? The authors would not agree. It is not possible, of course, to generalise from this one sample. It points to the need for fuller investigation on the relation of the MTAI to the personality Questionnaire, comparable to that done for the MMPI in America (Gowan and Gowan, 1955; Moore, Cole 1957). Allen (1963), reporting some of the American research studies which have used the MTAI to advantage, suggests that good use could be made on this side of the Atlantic of a scale similar in scope to the MTAI but adapted to conditions here.

(iv) *Interests.*

The Kuder Preference Interest Scores seemed to divide into two clusters: high Outdoor-Art-Music-Social Service scores and low Clerical-Mechanical-Computational-Scientific-Persuasive scores. This was the pattern for all the colleges and for all the social groups. It was surprising to find the men lowest on outdoor, mechanical and scientific interests, and perhaps equally surprising to find them highest on literary and artistic interests, contrary to the general findings of recent research (Strong, 1943).

CORRELATION OF INTEREST SCORES WITH TEACHING MARKS.

The only significant correlations with the teachings marks were as follows: in college 1, Persuasive and Clerical interests correlated negatively, but significantly with the teaching mark; in college 4, Mechanical Interests correlated positively, but Musical Interests negatively; there were no correlations at all for colleges 2 or 3. In college 1, Mechanical and Scientific Interest scores correlated positively with Intelligence, while in college 2, Outdoor Interests correlated negatively. All other correlations were insignificant. This would indicate and confirm the findings of previous research that the concept of interests is of little prognostic value in the selection of students for training as teachers (Evans, 1957).

IV.—DISCUSSION.

The varying and rather contradictory correlations of the different test results with the teaching mark point to great diversity of opinion as to what constitutes a good teacher. The following factors were found to have some relation with the criteria in at least three of the colleges:

MTAI: Positive for all four colleges, significant for colleges 2 and 3.

Factor M: Positive for all Irish colleges, significant for college 1.

Factor Q2: Positive but insignificant for all Irish colleges.

Factor Q4: Positive but insignificant for all Irish colleges.

Factor N: Positive but insignificant for three colleges.

Factor I : Positive but insignificant for three colleges.

Factor L : Positive but insignificant for all women's colleges.

Factor Q1 : Negative but insignificant for all Irish colleges.

Factor O : Negative but insignificant for the women's colleges.

Factor G : Negative but insignificant for three colleges, but positive and significant for college 1.

Scientific Interests : Positive but insignificant for three colleges.

Literary Interest : Positive but insignificant for three colleges.

Social Service Interest : Positive but insignificant for three colleges.

Mechanical Interest : Positive but insignificant for three colleges.

Some slight variations were found common to Social Groups II and III : Intelligence, MTAI, Factors G and Q2, and Musical Interests correlated positively, but insignificantly, with the teaching mark ; and Factors H, F, Q1, O, Persuasive and Social Service Interests correlated negatively, but insignificantly with it.

These results are in essence the same as those found by other investigators (Evans, 1952, 1953 ; Warburton, *et al.*, 1963). It should be noted that the vital factor with which the test scores have been correlated in these investigations is the teaching mark. This mark is in a sense a prognosis, a forecast or prophecy, of how those in charge think the student will turn out as a teacher. All prophecies are, of course, somewhat hazardous. The apparent lack of significant relationship between this mark and the other dimensions of intelligence, personality, attitudes and interests may be a valuable commentary on the teaching mark itself. There is the possibility that the teaching mark, is an unreliable *statistical* measure. It is necessarily a subjective measure, and subjective type examinations, e.g., essays, have been shown to be less valid and reliable than the more objective and standardised type. The mark as used at present is an attempt to *sum up* in a *quantitative* statement the observers' *qualitative* assessment of another's ability. Some more exact measurement seems to be called for if satisfactory *statistical* correlations are to be obtained.

But there is also another possibility : assuming that the teaching mark is valid and reliable, or even moderately objective, the lack of correlation with test results would seem to indicate that the test results are not good forecasters, at least in these islands. Further research is obviously needed.

V.—REFERENCES.

- ALLEN, E. A. (1963). Professional training of teachers. *Educ. Res.*, 5, 200-215.
 BIRKINSHAW, M. (1955). *The Successful Teacher*. Hogarth Press.
 CATTELL, R. B. (1931). The assessment of teaching ability. *Brit. J. Educ. Psychol.*, 1, 48-72.
 CATTELL, R. B., and STICE, G. F. (1962). *Handbook for 16 P-F Questionnaire*. I.P.A.T. *Classification of Occupations*. H.M.S.O., 1960.
 COOK, W. W., LEEDS, C. H., CALLIS, R. (1951). *Manual for M.T.A.I.* New York : Psychological Corporation.
 EVANS, K. M. (1952). Study of attitudes towards teaching as a career. *Brit. J. Educ. Psychol.*, 22, 63-70.
 EVANS, K. M. (1953). Further study of attitudes towards teaching as a career. *Brit. J. Educ. Psychol.*, 23, 58-63.
 EVANS, K. M. (1957). Is the concept of interest of significance to success in a teacher training course? *Educ. Rev.*, 9, 205-211.
 EVANS, K. M. (1958). An examination of the M.T.A.I. *Brit. J. Educ. Psychol.*, 28, 253-257.

- EVANS, K. M. (1961). An annotated bibliography of British research on teaching and teaching ability. *Educ. Res.*, 4, 67-80.
- GOWAN, J. C., and GOWAN, M. S. (1955). A teacher prognosis scale for the M.M.P.I. *J. Educ. Res.*, 49, 1-12.
- HEIM, A. *Manuals for group tests of Intelligence AH4 and AH5*.
- MOORE, CLARKE, H., COLE, D. (1957). Relation of the M.M.P.I. to practice teaching ratings. *J. Educ. Res.*, 50, 111-116.
- STRONG, E. K., Jnr. (1943). *Vocational Interests of Men and Women*. Stanford Univ. Press.
- TARPEY, M. C. J. (1963). *The relative importance of intelligence in the selection of students in Irish Training Colleges*. Unpublished thesis for M. Psych.Sc. Degree, Univ. College, Dublin.
- TUDHOPE, W. B. (1944). Motives for choice of teaching as a career. *Brit. J. Educ. Psychol.*, 14, 129-144.
- VALENTINE, C. W. (1934). Reasons for choice of teaching profession by university students. *Brit. J. Educ. Psychol.*, 4, 237-259.
- WARBURTON, F. W., BUTCHER, H. J., FORREST, G. M. (1963). Predicting student performance in a University Education Dept. *Brit. J. Educ. Psychol.*, 33, 68-79.
- VERNON, P. E. (1953). The psychological traits of teachers. In *The Year Book of Education*. London: Evans Bros.

(Manuscript received 14th October, 1965.)

CHANGES IN RELIGIOUS ATTITUDE AND PRACTICES AMONG STUDENTS DURING UNIVERSITY DEGREE COURSES

By

G. W. PILKINGTON, PAMELA K. POPPLETON

AND

GILLIAN ROBERTSHAW

(University of Sheffield)

SUMMARY. A sample of 106 university students was given a religious attitude scale and a questionnaire concerning their religious activities and belief in their first year and then again in their third year. Significant decreases in response were found for four of the seven indices used. The movement away from religion appeared to be more common among women and Nonconformist and Roman Catholic students than among men and Church of England students. The extent to which the changes found in religious attitude and behaviour for the whole sample were the direct results of the students' experiences at the university cannot be determined by the present data. But a comparison made between the responses of the 106 university students and those of eighty-three students in their third year at training college suggests that undergoing a university education is, in itself, unlikely to be important, except for women and members of devout religious denominations, as mentioned above. The hypothesis that students reading scientific subjects at the university are more likely to decline in their religious activities and attitude than are students reading other subjects was not conclusively confirmed by the findings reported.

I.—INTRODUCTION.

DURING recent years there has been considerable discussion about the religious beliefs and practices of university students in this country. In particular, many have been concerned with the effects, if any, of different kinds of university courses upon the students' religious attitudes and activities. Unfortunately, much of this interesting speculation has been conducted informally and impressionistically on the basis of what casual observation, random interviewing and intuition seem to suggest the students are doing (Furlong, 1961; Stacey, 1962). It is true that some investigators have made incidental studies of students' religious practices when they were mainly concerned with other issues (Thoday, 1957; Zweig, 1963), but the amount of such work compares poorly with the numerous empirical enquiries into students' religious behaviour that have been carried out in America (Argyle, 1958, pp. 42-46). Moreover, in this country there has been a number of interesting studies of the nature and development of religious ideas and activities among children and adolescents (Moreton, 1944; Hilliard, 1959; Goldman, 1962; Wright, 1962; Hyde, 1963). Attempts to extend such work into early adulthood, and more especially efforts to assess the effects of different sorts of higher education on the students' religious attitudes and practices, would seem, therefore, to be of some interest.

Two previous papers by the writers (1962, 1963) report an investigation into factors with which differences in the religious attitudes and activities of university students may be associated, notably, year of study, course of study being followed, age, denominational upbringing or affiliation and sex. The

method then used, however, was cross-sectional, that is, changes during the years at university were assessed by comparing the responses of samples of students from different year groups. The present paper is an account of an effort made (Robertshaw, 1963) to study changes in religious attitudes and practices among students over their years at the university by the more satisfactory and realistic longitudinal or 'follow-up' method.

II.—METHOD OF INVESTIGATION.

In February, 1961, a proportionate, stratified random sample of 500 was taken from the 3,000 full-time students there then were at the University of Sheffield. This sample included 141 students then in their first year at the university. Of these, only 112 survived into their third year. In December, 1962, these students were circulated with (a) an attitude scale for measuring religious attitudes in a university population, and (b) a questionnaire about typical religious practices and belief. These were exactly the same in form as those they had completed twenty-one months earlier. The full scale and questionnaire, together with details concerning the construction and scoring of the scale and evidence of its reliability, validity and unidimensionality are given by the authors (1963, 1964). As before, a covering explanatory letter and a stamped, addressed envelope for reply were sent to each student. After two follow-up letters, 106 students returned properly completed scales and questionnaires, which gives the satisfactory response rate of 95 per cent. The replies of the first year students to the 1961 survey were then reanalysed so as to exclude both the twenty-nine students who had since left the university and the six students who did not complete the forms a second time. In this way, the replies of the same students to the same attitude scale and questionnaire at two different stages of their university careers could be directly compared.

Finally, copies of the scale and questionnaire were given to a group of eighty-three third-year students at a Sheffield Training College. It was thought that such a group of students would be likely to be similar to the university students on such variables as age and educational background and, therefore, that a comparison of their responses with those of the third year university students might prove interesting. Unfortunately, the proportions of the two sexes within the two groups were very different, the training college group containing a much higher percentage of women than the university group. As sex is known to be an important determinant of religious attitude and behaviour, comparisons were, therefore, possible only within the sexes.

III.—STATEMENT OF RESULTS.

(a) *Changes within the whole university sample.*

Table 1 shows the percentages of students reporting at the two different dates the six different types of religious activity and belief with which the questionnaire deals. The significance test used was the McNemar test for the significance of changes, which Siegel (1956, p. 63) recommends as particularly suitable for data of the present kind. The significance of changes by the students in their attitudes scale scores are also shown in Table 1. The test used was the Wilcoxon matched pairs signed ranks test (Siegel, 1956, p. 75). This showed that the number and size of decreases in scale scores since 1961 was significantly greater than the number and size of increases, even though the median scale scores at the two dates were very similar.

TABLE 1

COMPARISON OF UNIVERSITY STUDENTS' RESPONSES IN FIRST YEAR AND THIRD YEAR.
(N=106)

Religious Activity	Percentage, March, 1961	Percentage, December, 1962	Significance
Hold Religious Belief	76	65	$p < .05$
Say Private Prayers	63	58	N.S.
Attended Church	48	42	N.S.
Active Church Members	42	30	$p < .005$
Say Daily Prayers	32	31	N.S.
Member of Student Religious Society	22	15	$p < .05$
Changes in Attitude Scale Scores			$p < .015$

(b) Changes and responses within the sexes.

Tables 2 and 3 show the percentages of male and female university students reporting at the two dates six different kinds of religious activity and belief. Similar comparison data are shown for the sample of third-year training college students. The significance of changes by the university students in their attitude scale scores are also included. The significance tests used for assessing changes over time within the questionnaire and scale were the same as those mentioned above for the whole university sample. In the last column of the two tables are shown the result of testing for significance differences between the responses of the third year university and the third-year training college students, chi-square being used for the questionnaire findings and the median extension test for the scale scores.

A further result was that in March, 1961, the women university students were significantly more religious than the men, both on all six indices of religious practice and belief and in their median scale scores. But by December, 1962, the women were significantly more religious than the men only in saying private prayers; nor were the men's and women's median scale scores significantly different at the later date. In contrast, within the sample of third-year training college students, there were significant sex differences in terms both of median scale scores and of religious belief, church attendance and active church membership.

(c) Changes within Faculties.

The changes found for the whole sample in the six categories of religious activity (see Table 1) were confined mainly to students within the Faculties of Arts and of Pure Science, those observed within the Faculties of Medicine and Applied Science being negligible. But none of the changes within the Faculty of Pure Science amounted to a significant decline and only one within the Faculty of Arts, that of active church membership ($p < .05$). Similar findings resulted from an analysis of the changes in attitude scale scores within

TABLE 2

COMPARISON OF MALE STUDENTS' RESPONSES IN FIRST YEAR AND THIRD YEAR AT UNIVERSITY AND IN THIRD YEAR AT TRAINING COLLEGE.

Religious Activity	University Students (N=79)			Training College Students (N=26)	
	Percentage March, 1961	Percentage December, 1962	Significance	Percentage	Significance
Hold Religious Belief ..	65	62	N.S.	73	N.S.
Say private prayers ..	52	49	N.S.	77	$p < .05$
Attended Church	38	37	N.S.	42	N.S.
Active Church Members	32	27	N.S.	15	N.S.
Say daily prayers	25	25	N.S.	31	N.S.
Member of Student Religious Society	13	10	N.S.	8	N.S.
Changes in Attitude Scale Scores			$p < .05$	Difference in Median scale scores.	N.S.

TABLE 3

COMPARISON OF FEMALE STUDENTS' RESPONSES IN FIRST YEAR AND THIRD YEAR AT UNIVERSITY AND IN THIRD YEAR AT TRAINING COLLEGE.

Religious Activity	University Students (N=27)			Training College Students (N=57)	
	Percentage March, 1961	Percentage December, 1962	Significance	Percentage	Significance
Hold Religious Belief ..	92	75	$p < .05$	95	$p < .01$
Say private prayers	89	81	N.S.	89	N.S.
Attended Church	70	56	N.S.	65	N.S.
Active Church Members	67	41	$p < .02$	46	N.S.
Say daily prayers	52	44	N.S.	32	N.S.
Members of Student Religious Society	48	25	$p < .05$	12	N.S.
Changes in Attitude Scale Scores			N.S.	Difference in Median scale scores.	N.S.

the faculties. Use of the Wilcoxon test showed that the number and size of decreases in scale scores was significantly greater than the number and size of increases only within the Faculty of Arts ($p < .05$). For the interpretation of these findings, it should be borne in mind that the proportions of men and women students within the Faculties of Arts, Pure Science and Medicine were approximately $1\frac{1}{2} : 1$, $3 : 1$ and $4 : 1$, and that there were no women students at all within the Faculty of Applied Science.

(d) *Changes within denominations.*

All except three of the 106 students claimed either membership of, or having been brought up within, a religious denomination. Of these, seventy-one were Church of England, twenty-three Nonconformists and nine Roman Catholics. Within the Church of England, the proportion of students saying private prayers declined significantly ($p < .05$), and within the Nonconformists the proportions reporting active church membership and membership of a student religious society did likewise ($p < .05$). No significant decreases in attitude scale scores were revealed for either denomination. The Roman Catholic sample size was too small for the significance of their changes in response to the questionnaire items to be assessed; but significant decreases in their attitude scale scores occurred ($p < .01$). Finally, changes in interdenominational differences over the two years were noted. In March, 1961, the Roman Catholic students were considerably more devout than the Church of England students, showing a significantly higher proportion of positive response to four of the six questionnaire indices ($p < .02$). The Nonconformist students also had a significantly higher proportion of positive response than the Church of England students, but only to one questionnaire item ($p < .05$); but they were less religious than the Roman Catholic students, their median attitude scale score being significantly lower than that of the Catholics ($p < .05$). By December, 1962, these differences between the denominations were much reduced. The general 'religious rank order' of the three denominations—Roman Catholic, Nonconformist and Church of England—was still preserved, but now the Roman Catholic students were significantly more religious than the Church of England and Nonconformist students on only two and one of the questionnaire indices, respectively. Moreover, at the later date, there were no significant differences between any of the denominations' median attitude scale scores. Further details are given by Robertshaw (1963). A final point relevant to the interpretation of these data is that the differences in the proportions of the two sexes between the three denominations were slight and not statistically significant.

IV.—DISCUSSION OF RESULTS.

Argyle (1958) has summarised a number of investigations in America which seem to show that religious activities become less frequent and religious beliefs more 'liberal' among students during their early years at college. The authors' cross-sectional study of 463 Sheffield students in 1961 gave some degree of support to the view that this tendency might also be true of students at British universities, at least those within the Faculties of Arts and Pure Science. The longitudinal data reported above, however, confirm this tendency more convincingly. Table 1 shows that during their two years at the university the 106 students declined significantly on four out of seven indices of religious activity and attitudes. But it does not follow from this that a university education, in itself and of whatever kind, produces this decline. For other factors concomitant with attendance at the university might be responsible. One such factor is age.

Argyle (1958) mentions several British and American studies, which show a decrease in many different kinds of religious activity between the ages of 18 and 30; similarly, one of the writers' earlier investigations (1962) revealed a tendency for older university students to be significantly less religious on four of the questionnaire indices than were younger students. Without a control group from the general population matched with the university students for such important variables as age and similarity of previous educational background, no firm conclusion can be drawn from the present study concerning the extent to which the changes reported are the result of the students' experiences at the university. The sample of training college students could have been regarded as such a control group, if questionnaire and attitude scale data from them in their first year had been available. But, as matters stand, comparison of their scores in the third year with those of similar university students (as shown in Tables 2 and 3 above) shows very few significant differences. This could be interpreted as suggesting that a university education, as such, has little effect on the students' religious activities and attitudes. On the other hand, particular kinds of subjects studied within the university may make a difference. Robertshaw (1963) interviewed the twenty-four students whose attitude scale scores had most increased or decreased over the two years; the interview schedule used was designed to elicit from the students those factors which had, in their opinion, primarily brought about the changes in their religious attitudes. As might be expected, for both groups of students the influence of friends and of discussions emerged as the single, most important determinant of change. But the effect of subject of study at the university was mentioned by none of those who had increased their scale scores, whereas it was thought to be a major influence by over half those who had decreased their scale scores; the subject of study mentioned was always a scientific one.

Within the sample, the most interesting results are those which show that the women university students have, over the two years, moved away from religion more than have the men. Tables 2 and 3 show that the men have declined significantly on only one of the seven indices of religious activity and attitude, but that the women have done so on three. Moreover, by the third year, the women university students are significantly more religious than the men on only one index of religious behaviour, whereas in the first year all seven indices revealed significant sex differences. This finding may be contrasted with that for the third year training college students, among whom four of the indices showed significant differences between men and women. It is tempting to conclude that most of the decrease in religious activity shown for the whole sample of university students in Table 1 is attributable to the women rather than to the men students.

The marked reduction found in differences between the sexes in religious behaviour and attitude over years at the university is surprising and unexpected. Argyle (1958) has shown that one of the most firmly established generalisations about religious behaviour is that women are more religious than men on many different criteria; and the writers have shown that this tendency certainly holds true of their earlier sample of students (1962, 1963). Clearly, however, no explanation of the relatively greater decline by the women university students in their religious practices can be given on the basis of the present data, although a possible hypothesis is that this trend may be accounted for as a process of increasing conformity to the lower group religious norm of a predominantly male community. But what seems most to be needed at present are efforts to determine whether this tendency is true of larger groups of women students at a number of different universities.

Analysis of the changes within Faculties revealed that only students within the Faculty of Arts showed any significant decreases in religious practice and attitude, and this is very probably explained by the fact that this Faculty contained the largest proportion of women students. Apart, perhaps, from the interview findings, no firm evidence emerged to support the tendency reported in the 1961 cross-sectional investigation for students within the Faculty of Pure Science to become less religious over their years at the university. Certainly, all categories of religious activity and attitude within this Faculty showed a decline, but none of these was statistically significant, so that the most that can be said is that these data neither confirm nor refute the earlier findings. The interview data referred to earlier can be regarded as, at most, suggestive.

Many American studies, for example, those by Allport, Gillespie and Young (1948), Telford (1950), Brown and Lowe (1951) and Gilliland (1953), have shown that Roman Catholic students tend to be more religious on many different criteria than do students belonging to other denominations. The writers in their earlier papers referred to above confirmed this tendency within their sample; they also found that Nonconformist students were significantly more religious than Church of England students, but significantly less religious than the Roman Catholics. A similar relationship between the denominations was apparent among the present sample of 106 university students during their first year, but was less evident, although still perceptible, by their third year. A possible explanation of this trend may be that students within the two most religious denominations, the Roman Catholics and the Nonconformists, declined in their religious activities and attitude more than did the Church of England students. Some tentative support for this hypothesis is given by the findings presented above in Section III (d). Again, there is perhaps, a process taking place of increasing conformity to the lower group religious norm of the majority of the students at the university. But the sample sizes of the two more religious denominations are very small, especially for the Roman Catholics, so that no certain conclusions can be drawn.

V.—CONCLUSIONS.

The findings presented in this paper suggest that, during the early years of a university course, there may be a significant decline in the students' religious attitudes and activities; that this decline is likely to be more common among women students than among men students; and that students brought up within the Roman Catholic and Nonconformist denominations might show this tendency more markedly than students brought up within the Church of England. In general, it would seem that initial differences between particular categories of students in their religious activities and attitude decrease during their years at the university, a tendency which is possibly explainable as a process of increasing conformity to a group norm which is neither extremely favourable nor extremely unfavourable to religion. Evidence for any different effects of certain sorts of university courses on the students' religious behaviour was inconclusive. In particular, the present data, although consistent with the authors' earlier tentative finding that a course within the Faculty of Pure Science may make the students less religious than previously, do not conclusively provide evidence for such a tendency. Similarly, the results reported are not sufficient to determine whether a university education in itself results in a movement by the students away from religion, although the comparison findings from the training college students do seem to point against this hypothesis. What, therefore, is now most needed are further attempts to replicate the

findings reported above amongst larger samples of students from a representative range of universities and efforts to check any changes found against suitable control groups.

VI.—REFERENCES.

- ALLPORT, G. W., GILLESPIE, J. M., and YOUNG, J. (1948). The religion of the post-war college student. *J. Psychol.*, 25, 3-33.
- ARGYLE, M. (1958). *Religious Behaviour*. London: Routledge and Kegan Paul.
- BROWN, D. G., and LOWE, W. L. (1951). Religious beliefs and personality characteristics of college students. *J. Soc. Psychol.*, 33, 103-129.
- FURLONG, M. (1961). Religion in the universities. *Sunday Times*, 11th June, 28.
- GILLILAND, A. R. (1953). Changes in religious beliefs of college students. *J. Soc. Psychol.*, 37, 113-116.
- GOLDMANN, R. J. (1962). *Some Aspects of the Development of Religious Thinking in Childhood and Adolescence*. Monograph of Institute of Education, University of Reading.
- HILLIARD, F. H. (1959). The influence of religious education upon the development of children's moral ideas. *Brit. J. Educ. Psychol.*, 29, 50-59.
- HYDE, K. E. (1963). Religious concepts and religious attitudes, I and II. *Ed. Rev.*, 15, 132-141, 217-227.
- MORETON, F. E. (1944). Attitudes to religion among adolescents and adults. *Brit. J. Educ. Psychol.*, 14, 69-79.
- PILKINGTON, G. W., and POPPLETON, P. K. (1962). Recent trends in religious belief and practice among university students. Mimeographed paper read to the Annual Conference of the Education Section of the B.P.S., London, September, 1962.
- POPPLETON, P. K., and PILKINGTON, G. W. (1963). The measurement of religious attitudes in a university population. *Brit. J. Soc. Clin. Psychol.*, 2, 20-36.
- POPPLETON, P. K., and PILKINGTON, G. W. (1964). A comparison of four methods of scoring an attitude scale in relation to its reliability and validity. *Brit. J. Soc. Clin. Psychol.*, 3, 36-39.
- ROBERTSHAW, G. (1963). *An Empirical Enquiry into Students' Religious Attitudes and Practices*. University of Sheffield: B.A. degree dissertation.
- SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioural Sciences*. New York: McGraw-Hill.
- STACEY, N. (1962). Youth, faith and the church. *Sunday Times*, 10th June, 21.
- TELFORD, C. W. (1950). A study of religious attitudes. *J. Soc. Psychol.*, 31, 217-230.
- THODAY, D. (1957). Halls of residence. *Univ. Quart.*, 12, 45-56.
- WRIGHT, D. S. (1962). A study of religious belief in sixth form boys. *The University of Leeds Institute of Education Researches and Studies*, 24, 1-8.
- ZWEIG, F. (1963). *The Student in the Age of Anxiety*. London: Heinemann.

(Manuscript received 1st June, 1964.)

THE APPLICATION OF PIAGET'S SCHEMA OF OPERATIONAL THINKING TO RELIGIOUS STORY DATA BY MEANS OF THE GUTTMAN SCALOGRAM

By R. J. GOLDMAN

(*University of Reading*)

SUMMARY. A research into children's capacity to understand religious concepts was designed on a clinical interview basis. A standardised procedure was devised to elicit responses to religious pictures and stories and administered to a stratified random sample of 200 pupils with ages ranging from 6 : 1 to 17 : 11. Some forty responses for each child were evaluated on the criteria of Piaget's schema of operational thinking. Scoring reliability was safeguarded in the Piagetian evaluation by the use of independent judges familiar with the criteria and a suitable level of correlation between the investigator's and the independent judges' scores. Assessment of the validity of Piaget's schema was made by arranging scores in order of chronological age, mental age and total score and the use of the Guttman Scalogram. In all five items selected suitable reliabilities were found and the sequences of Piaget's schema of operational thinking were seen to be valid when applied to religious story data. Several discrepancies of age levels and general results are discussed, together with the implications for a theory of religious cognition.

I.—INTRODUCTION.

In an earlier article in this *Journal*, E. A. Peel (1959) outlined several experiments at Birmingham to test the validity of Piaget's schemata of mental development of children. The areas of haptic perception, spatial relationships in drawing, logical judgments and moral judgments were investigated. The present writer set out to investigate the field of religious cognition, including pupils' theological judgments, utilising much of the experimental technique outlined by Peel. The area investigated in religious cognition, however, was fairly homogenous: the sample of children was 200 in number and the age range more extensive. The purposes of this investigation were:

- (1) To ascertain a variety of religious concepts which are central to any understanding of religious stories and activities from the age of 6 years to the end of the 17th year.
- (2) To discover what sequences, if any, exist, by which the pupils' levels of understanding progress.
- (3) To evaluate the logical processes used by children and adolescents at varying stages of development, taking Piaget's schema of the development of operational thinking as a comparative guide.
- (4) To examine what factors may influence religious thinking, especially the influence of home, the religious affiliation of child and parents, chronological age and mental ability.

II.—THE DESIGN OF THE RESEARCH.

Preliminary discussions on biblical and picture material were held with children aged 5 to 17 years, and from these a story and picture religious thinking test was constructed of five simple pen and ink drawings and eight biblical stories. Pre-testing reduced these to three drawings and three biblical stories,

selected as those which invoked the most widely used religious concepts in children's thinking. The drawings, designed as a structured projective technique, are not the concern of this article and we will concentrate upon the story results. The final selection of stories was Moses and the Burning Bush, The Crossing of the Red Sea, and the Temptations of Jesus, all recommended at practically all ages by some Agreed Syllabuses. The stories were played on a tape recorder to individual children and their responses to questions after each story were noted down verbatim. Twenty children, used as a pilot sample, were tested, revisions were made and the final form of the test administered to a sample of two hundred children.

Sampling was upon a stratified random basis and procedures designed to ensure as typical a distribution as possible in terms of age, sex, ability, home background and religious (or non-religious) affiliation. Pre-testing had revealed that children below the age of six were not responsive to the test situation and the sample, therefore, began at 6, with ten boys and ten girls in each year, with a median age of 6 months between years. Table 1 shows distribution by ability.

TABLE 1

DISTRIBUTION OF SAMPLE IN YEAR AGE GROUPS AND MEAN AVERAGE INTELLIGENCE SCORES.

Year	I.Q. Range	S.D.	Mean I.Q. Boys	Mean I.Q. Girls	Total I.Q.
6	77—140	22.50	103.7	105.3	104.5
7	79—132	20.62	103.9	107.5	105.7
8	79—146	19.34	103.6	109.4	106.5
9	76—140	21.86	109.2	106.8	108.0
10	78—144	19.02	104.0	107.2	105.7
11	77—140	20.88	108.4	107.9	108.1
12	78—140	19.30	104.6	107.9	106.25
13	76—140	19.30	107.1	104.1	105.9
14	76—140	19.30	104.9	106.3	105.6
15	91—140	—	118.55	119.8	119.1
16	104—140	—	—	—	—
17	110—137	—	—	—	—
Total I.Q. incl. 15—17			106.77	108.24	107.5
Total I.Q., exclusive 15—17			105.46	106.95	106.2

The 15—17 year-old age range constitutes the main sampling problem. The variable of 'after leaving school influence' would predominate in a selection of adolescents representing the normal curve of the distribution of intelligence, since practically all those of lowest ability left school at the first opportunity. The lowest I.Q. noted in this age range remaining at school was 91, a few at 103 to 105 and the rest 'superior.' This, however, is to be expected since mostly those capable of profiting from Further Education in an academic sense tend to stay on. It must be made clear, therefore, that the 15+ sample is included but does not meet the requirement of 'typicality.' There is a slightly superior mean average score of girls' intelligence quotients over boys'.

An F test was applied to see whether, taking the widest difference of standard deviations between the age groups, any significant difference could be found. The widest difference between the standard deviations was in year 6

(S.D. 22.50) and year 10 (S.D. 19.02). *F* was found to be 1.39 which, at 19 and 19 degrees of freedom is not significant at the 5 per cent. level. There is, therefore, no statistically significant evidence against the hypothesis that the variability of I.Q. is the same in all age groups.

Taking the average within-groups-S.D. as 20, the standard error for comparing any two age groups means is 6.33. To be significant, the difference between any two group means would have to exceed $2 \times \text{S.E.}$, or at least 12.66 I.Q. points, it is evident that the mean I.Qs. do not differ significantly. Differences, therefore, in I.Qs. are unlikely to introduce bias between age group comparisons. The exception is the 15-17 year-old sample.

Since we are not using a completely random sample, most normal statistical assumptions are not applicable to our data, since various variables may operate in each age group separately. In the method of evaluating the data, namely, the use of the Guttman scalogram technique, this objection is not valid, since comparisons are made over the whole span of age and ability and not from year to year.

The typicality of religious affiliation (and non-affiliation) was attempted, as can be seen in Table 2.

TABLE 2
THE RELIGIOUS AFFILIATION OF PUPILS IN THE SAMPLE.

Year	6	7	8	9	10	11	12	13	14	15-17	Totals
Church of England	9	9	9	9	9	9	8	9	9	10	90
Free	4	4	4	4	4	4	5	5	5	3	42
Gospel	1	1	2	2	1	2	2	1	1	1	14
Nothing	6	6	5	5	6	5	5	5	5	6	54
Total											200

It was clearly necessary to provide a controlled proportion in each year age group of the various church and non-church allegiances possible. It was also evident that of attenders at Sunday School or Church the largest group would be Church of England, the next largest the Free Churches and the smallest would be the Gospel sects, such as Pentecostal, Adventist and Salvationist Churches. Roman Catholics are not included since they would normally be in separate schools or be withdrawn from religious instruction. The crude numbers of attenders would appear to be in excess of all estimates (see Argyle and Hyde) of the proportion of children connected with the churches. This sample, for example, suggests 20 per cent. of children are connected with the Free Churches when Hyde estimates approximately 15 per cent. in 1954.

These crude figures, however, conceal wide difference of frequency of attendance. 'Nothing,' is the appellation given to a child who goes nowhere at present, has not been to Church or Sunday School for over a year and who has neither parent a church-goer, a rather stringent requirement for this category. Figures for attenders may conceal the fact that a child calls himself C. of E., but only goes at Easter and is, to all purposes a very marginal member.

A major problem in attempting to control the variable of Church or Sunday School attendance is the fact that from 11 years onwards there is a recognisable

trend of declining attendance. If this decline were to be reproduced in the sample, it would introduce variable factors, hence the attendance proportions have been kept as constant as possible. No attempt will be made to generalise the results in terms of denominations since the numbers of each denomination are too small as separate samples and too many other variables are involved.

Since occupational and social factors in the family may introduce a bias in the sample from year to year, fathers' occupations were reduced to scores based upon the 1951 Census "Classification of Occupations" as 1 to 5. The median score for each year age group fathers' occupations was 3.0, the total mean average score for all ages being 2.975.

Sampling of schools and classes was varied to ensure that bias of particular schools and teachers was not introduced. All schools were State schools in urban areas in the West Midlands and Southern England.

III.—FIVE QUESTIONS SCORED ON PIAGET'S OPERATIONAL THINKING SCHEMATA.

Five questions were chosen to discern possible logical processes on the pupil's religious thinking and evaluated on an operational scoring scale. These were:

- (1) "Why was Moses afraid to look at God?"
- (2) "Why was the ground, on which Moses stood, holy?"
- (3) "How would you explain the bush burning, but not being burnt?"
- (4) "How would you explain the dividing of the waters of the Red Sea?"
- (5) "Why wouldn't Jesus turn the stone into bread?"

The investigator scored them on a seven-point scale (0-6) representing the structure of Piaget's criteria of pre-operational, and operational thinking. The answers for each question were then sent separately to five independent judges to be scored on the same seven-point scale. Neither the investigator's scores nor the identity, sex, age, or other details of the pupils were disclosed to the independent judges in order to avoid, on the one hand, biased marking, and on the other, 'halo effect.' All the independent judges chosen were acquainted with Piaget's terminology. Scores were allocated on the following basis:

Pre-operational answers	Score 1
Intermediate between pre-operational and concrete operational answers	Score 2
Concrete operational answers	Score 3
Intermediate between concrete and formal operational answers	Score 4
Formal operational answers	Score 5
Advanced formal operational answers	Score 6

The independent judges were specifically asked to evaluate on the Piagetian scale and not to evaluate the theological worth of the answers.

When the independent judges' score sheets were returned, a Pearson product-moment correlation-coefficient was computed between the investigator's scoring and the independent judge's scoring of 200 answers to the one question.

The standard deviation of the differences between examiners is a useful measure of the reproducibility of marks. It allows for the average differences in the number of marks between examiners. It can be interpreted in two ways: First, 95 per cent. of the time two examiners should agree about the mark to be given to a particular response to be within ± 2 of these standard deviations.

Second, if the two examiners are equally reliable and use an equal range of marks, then 95 per cent. of the time, a single examiner will assess the difference in marks for the children within ± 2 of these standard deviations of the 'true' difference in marks between the children. Thus, in item 4 where the S.D. of differences in marks is highest at 0.823, 95 per cent. of the time the investigator and the independent judge should agree to within about $\pm 2 \times 0.823$, that is 1.646 marks.

TABLE 3
INTER-SCORER RELIABILITY COEFFICIENTS.

Item	Correlation Coefficients	Agreed	Differences			S.D. of Differences
			1 pt.	2 pt.	3 pt.	
1	.84	119	69	12	0	0.678
2	.74	117	67	15	1	0.775
3	.79	138	55	5	2	0.619
4	.75	117	63	17	3	0.823
5	.86	144	48	8	0	0.611

This is also brought out in the number of agreements and disagreements set out in detail in Table 3. When we consider that many of the discrepancies in scoring are repeated many times, the results appear to be satisfactory. The overall results appear to indicate that the items were capable of being objectively scored.

Where there were discrepancies of scoring between the investigator and the independent judge on any pupil's response, the 'average' between them was allocated as the final score. Where only a 1 point difference was evident, then the higher score was allocated. This leads to a slight inflation of scoring upwards, but its effect is very small.

IV.—THE GUTTMAN SCALING.

The sample of 200 subjects and the items of the test were so arranged that scale analysis could be applied to the data. The technique of scale analysis can be used with qualitative data and has been applied, in the main, to the measurement of attitudes and opinions. Peel (1959) suggests that Guttman's scalogram analysis can be used to express the discrimination of the Piaget schema in the case of logical judgments, as a coefficient of reproducibility taking chronological age, mental age and total level as criteria. A single item may be evaluated in this way, or any number of items may be combined, to see whether they form a single scaleable universe. The technique has been described by Guttman (1944 and 1947) and Stouffer and Guttman (1950).

Once scores are allocated on a seven-point scale (0-6) they may be set out in various ways. One way is to arrange the scores in seven category columns in order of increasing or decreasing rank. This is used, for example, by Coltham (1960) in evaluating children's understanding of historical terms. The scores can also be arranged in order of chronological age, mental age, or in order of total scores (total rank order).

We would expect that if levels of insight are related to mental age, the scores would move gradually from left to right, i.e., from lower scores to higher scores, so that the familiar parallelogram of scores would occur. With this ideal scoring, the reproducibility can be said to be 100 per cent. Usually, the reproducibility is far from perfect. In other words, some children of low mental age score higher than some of those of superior mental age, and the ideal parallelogram does not occur.

To improve reproducibility expressed as a coefficient, it is legitimate to combine categories of scores. Guttman, (1947), elaborates the reasons for this:

"It has seldom been found that an item with four or five categories is regarded as distinct. One reason for this is the verbal habits of people... (they may) have essentially the same position in the basic continuum, but differ on an extraneous factor of verbal habits. By combining categories, minor extraneous variables of this kind can be minimised." (p. 256).

This proved to be necessary with the five selected items and the scores of four of them were combined in the following manner:

Scores 0, 1 and 2 were combined to score 1.

Scores 3 and 4 were combined to score 2.

Scores 5 and 6 were combined to score 3.

The one exception was item 2, where score 0, 1 and 2 were combined, 4, 5 and 6 were combined, and score 3 stood alone as a final intermediate category.

In calculating coefficients of reproducibility and evaluating their significance several rules were observed, based upon Guttman's criteria. These were:

- (i) "No category shall have more error in it than non-error" (Guttman, 1947, p. 260).
- (ii) "To avoid spuriously high coefficients of reproducibility due to very high frequency in one category, only those co-efficients are regarded as satisfactory which have the best distribution of marginal frequencies." (Guttman, 1947, p. 260).
- (iii) When less than ten items are used and combined, "it may not be safe to assume that the universe is scalable if all the items must be dichotomised in order to obtain high reproducibility; at least some would be retainable in trichotomised form in order to make the inference plausible" (Stouffer, 1950, p. 117).
- (iv) "Where items are combined, and some are trichotomised and some dichotomised, some dichotomies should have marginals close to 50 : 50 to avoid spuriously high reproducibility."

V.—THE RESULTS.

For the five questions, scored and scaled on the Guttman technique, coefficients of reproducibility were calculated. For typical responses, see the Appendix. Guttman, in scaling attitude questionnaires, says Peel (1959) "insisted upon a minimum of 85 per cent. Guttman's condition is very stringent and not often met, and in this work (scaling logical judgments) we might say that 75 per cent. is a reasonable figure to take." The view is taken here that any coefficient of reproducibility which approaches or meets Peel's 75 per cent. requirement is sufficient to justify the existence of sequences in thinking, based upon the scoring criteria given.

The 75 per cent. requirement is particularly appropriate in this investigation, since no attempt has been made to move the cutting lines up and down, and so ignoring the marginal frequencies, as is permissible in attitude scaling. The coefficients of reproducibility reported are obtained from a very rigorous application of scaling techniques and tend, therefore, to be on the conservative side.

Table 4 shows coefficients of reproducibility achieved for individual items and combined on the criteria of chronological age, mental age and total scores.

TABLE 4

COEFFICIENTS OF REPRODUCIBILITY ON CRITERIA OF CHRONOLOGICAL AND MENTAL AGE AND ON TOTAL SCORE.

Criterion	Combined Categories 1 2 3			Total Errors	% Age Errors	% Coeff. of Reprod.	Cutting Lines
Chron. Age :							
Item 1.....	11	32	23	66	33.0	67.0	7 : 6 and 13 : 6
2.....	15	42	30	87	43.5	56.5	7 : 10 and 12 : 3
3.....	8	32	24	64	32.0	68.0	6 : 8 and 12 : 10
4.....	17	34	18	69	34.5	65.5	7 : 6 and 13 : 6
5.....	16	40	25	81	40.5	59.5	8 : 6 and 13 : 6
Totals				367	36.7	63.3	
Mental Age :							
Item 1.....	11	24	14	49	24.5	75.5	7 : 11 and 14 : 2
2.....	12	38	27	77	38.5	61.5	8 : 2 and 12 : 8
3.....	7	30	23	60	30.0	70.0	6 : 6 and 13 : 5
4.....	14	31	17	62	31.0	69.0	7 : 10 and 14 : 2
5.....	18	35	17	70	35.0	65.0	8 : 10 and 14 : 2
Totals				318	31.8	68.2	
Total Scores :							
Item 1.....	6	13	7	26	13.0	87.0	
2.....	10	30	20	60	30.0	70.0	
3.....	6	27	21	54	27.0	73.0	
4.....	11	19	8	38	19.0	81.0	
5.....	10	21	11	42	21.0	79.0	
Totals				220	22.0	78.0	

Since there is a variety of data, a different level of word-understanding, and other differences in each story, it seems legitimate to take the highest score in all five items and scale them against the usual criteria. The same trends, with slightly higher reproducibility are evident in Table 5.

The mean average level of operational thinking is a slightly better indicator than a pupil's highest score, and can be seen in Table 6. No reproducibility is calculated against the criterion of total score for mean average score since by the nature of the computation, the result is bound to be 100 per cent.

TABLE 5

COEFFICIENTS OF REPRODUCIBILITY WHEN HIGHEST SCORES ARE SCALED.

Criterion	Combined Categories			Total Error	% Age Error	% Coeff. of Reprod.
	1	2	3			
Frequency	37	65	98			
Chronological Age	11	33	25	69	34.5	65.6
Mental Age	11	24	18	53	26.5	73.5
Total Score	8	19	12	39	19.5	80.5

TABLE 6

COEFFICIENTS OF REPRODUCIBILITY WHEN MEAN AVERAGE SCORES ARE SCALED.

Criterion	Combined Categories			Total Error	% Age Error	% Coeff. of Reprod.
	1	2	3			
Frequency	29	61	110			
Chronological Age	10	25	18	53	26.5	74.5
Mental Age	7	19	13	39	19.5	80.5

The development of sequences of operational thinking is least marked by mental age and most marked by total score when the best arrangements of tied series is made. This validates Piaget's general structure or sequence of thinking, in three stages, as applied to the pupil's thinking about religion.

While no age ranges can be firmly established as peculiar to any given phase in the sequence, the cutting lines in Table 4 indicate some interesting indications of boundaries of thought, where the majority of children appear to move into a higher type of thinking. The beginning of the phase of concrete operations tends to be on these five items, within the general age 'area' suggested by Piaget, but as with Lodwick (1958), there is a variation of each pupil, sometimes slight and sometimes large, from item to item. This appears to vary with the level of difficulty posed by the data used in each story. Taking the mental ages as more accurate indicators of boundaries, we may note that pupils tend to arrive at the concrete operational stage very early in item 3 (M.A. 6:6) where the burning bush miracle is discussed; late in item 1 (M.A. 7:11) and item 4 (M.A. 7:10) where fear of God and concepts of control of nature are concerned; and later still in item 2 (M.A. 8:2) and item 5 (M.A. 8:10) where concepts of 'holy' and ideas of temptation and messiahship are central. The very late arrival at concrete operations in the last story (M.A. 8:10) is indicative that such material may properly belong to the secondary school level.

A similar pattern is discernible where the stage of propositional thinking or formal operations is reached. Taking mental age recorded in item 2 (M.A. 12 : 8) is, however, somewhat misleading, since here a different combination of categories had to be made and scores 4, 5 and 6 (intermediate propositional thinking—score 4—and full propositional thinking—scores 5 and 6) were brought together. This means that M.A. 12 : 8 is only a boundary indicator between concrete operations and intermediate thinking, where propositional thinking is attempted, but is still limited by concrete elements and does not attain full formal operations. With this one exception we can see in item 3 (M.A. 13 : 5) that the earlier propositional thinking becomes evident ; and the other three items all reveal a boundary at about 14 : 2. Again, the level of difficulty offered by the data may be one hypothesis by which this may be explained.

These findings on the beginnings of propositional thinking would substantiate Peel's findings (1961) "that there is little evidence of children's capacity to set up possibilities to account for events in stories, as opposed to mere describing, before the age of 13 plus." Peel's further evidence, gathered in New Jersey, on story data, yields a general boundary of intermediate hypothetical thinking from a mean average age of about 13 : 1 which only yields to the highest level of hypothetical thinking between a mean average mental age of 14 : 8 to 15 : 0. The ages in his sample differ, as in ours, with the variety of story data involved.

The five questions assessed by psychologists on Piagetian criteria were also assessed by theologians on theological criteria. A product moment correlation coefficient produced the following results :

TABLE 7
CORRELATION BETWEEN OPERATIONAL AND THEOLOGICAL SCORES.

Item	r
1	0.803
2	0.797
3	0.776
4	0.787
5	0.891

In all cases the standard deviation of the differences is satisfactory. The two lowest correlation coefficients (items 3 and 4) both involve concepts of miracle where non-rational factors may exercise a stronger influence on thinking than in other questions.

VI.—CONCLUSIONS.

From the five items subjected to 'operational' thinking criteria and the results calculated by means of the Guttman scalogram technique, the following general conclusions can be made.

(1) Five rules were observed and all scaling resulted in suitable trichotomous sequences, both in the individual items and in the total scores. Few of the response categories (0-6) are independent since scalability was achieved only by combining categories to give three levels.

(2) An inspection of Table 4 will show that the threefold sequence is least marked by chronological age, and more marked by mental age. The best results occur when the criterion of total score is used, when the 75 per cent. reproducibility requirement is easily met.

(3) The same result is further evident with even higher levels of reproducibility when the highest scores and the mean average scores in the five items are scaled. Highest score is felt to be a most suitable indicator of the pupil's capacity or level of thinking and the threefold sequence is least marked by chronological age (C. of Rep. 65.5 per cent.), more marked by mental age (C. of Rep. 73.5 per cent.) and very marked by total score (C. of Rep. 80.5 per cent.).

(4) While Guttman's 85 per cent. criterion is not attained, results are sufficiently high to make the assumption that the five 'operational' items form a scale. Various variables appear to exist, such as age, ability, variations of motivation, familiarity and unfamiliarity with the data.

(5) The dominant variable appears to be progress in the grasp of the material, or in levels of understanding, as structured by the three stages of operational thinking. This overall result appears to validate Piaget's structure of pre-operational, concrete operational and formal operational thinking as applied to the realm of religious thinking.

(6) While the three-stage sequence appears to be supported, the ages tend to vary from one item to another considerably. Mental age appears to be a sounder indication of 'boundaries' of thinking than chronological age. As the data and level of difficulty varies from item to item, the greatest proportion of children tend to achieve concrete operations anywhere between M.A. 6:6 and 8:10, and propositional-hypothetical thinking or formal operations between M.A. 13:5 and M.A. 14:2. (One exception was noted as earlier, due to re-arrangement of categories). This latter boundary appears to substantiate Lodwick's (1958) and Peel's (1962) findings. These ages are only advanced tentatively, and as general 'areas' rather than fixed lines of age to be crossed. The variability of 'within pupil performance' is considerable.

(7) Correlations between the scoring of the same five operational items on psychological and theological criteria separately, are sufficiently high to assume that normal logical and rational processes occur in theological judgments, and that religious thinking can be developed by pupils only as far as their current levels of operational thought will allow.

VII.—IMPLICATIONS.

The implications of this and other related research for education are discussed elsewhere (Goldman, 1964 a and b) but the results of the five items subjected to a Guttman scaling on Piaget's operational thinking criteria yield some interesting implications for a theory of religious cognition.

First, the evidence appears to support a generalised theory of religious thinking as opposed to the concept of a religious faculty or instinct. Religious thinking appears to take place, within the context of belief in the supernatural, according to the same processes and methods of thinking as applied to other fields of experience. Religious thinking may, therefore, be defined in this generalised sense as thinking applied to the area of experience we designate as religious, rather than a specific faculty, drive or instinct. This is not to deny that religion is a 'natural' activity of man, but rather bases religious thought upon a more realistic theory.

Second, the age levels cited and interpreted as general areas when higher levels of thought become possible would seem to sustain the assumption that religious thinking is secondary in character, based upon the normal mental development of children in terms of the primary ingredients of sensation, perception and conceptual thinking. Since religious thinking is essentially an interpretation of experience, children must have lived long enough, have had sufficient primary mental stimulation and perceptual activity, and developed some capacity for propositional thought (Piaget's formal operational level) before the profound truths of religion can be understood at an insightful level.

Thirdly, the results further reinforce the problems posed by religious language, which is a secondary language, using metaphor, simile, analogy, and parable as its dominant forms. Until the original experiences of which biblical stories speak in these terms can be understood, the inferences drawn from them in describing religious truth will be distorted or misunderstood.

Finally, in terms of the Piagetian schema of pre-operational, concrete operational, and formal operational thinking, the limitations and the possibilities of these stages would appear to provide support for the evidence of this investigator (1964 b) that children pass through religious conceptual stages designated as pre-religious, sub-religious and religious. Whether the final stage is achieved or endures through adolescence seems to depend upon whether adolescents are able and willing to apply the higher levels of formal operations to religion.

VIII.—REFERENCES.

- ARGYLE, M. (1958). *Religious Behaviour*. London: Routledge and Kegan Paul.
- COTLHAM, J. B. (1960). *Junior School Children's Understanding of Some Terms Used in the Teaching of History*. Unpub. Ph.D. dissertation, University of Manchester.
- GOLDMAN, R. J. (1963). Children's Spiritual Development in *Studies in Education—First Years in School*. London: Evans' Bros.
- GOLDMAN, R. J. (1964 a). Researches in religious thinking. *Educ. Res.*, Vol. VI, 2, 139-155.
- GOLDMAN, R. J. (1964 b). *Religious Thinking from Childhood to Adolescence*. London: Routledge and Kegan Paul.
- GUTTMAN, L. (1947). The Cornell technique for scale and intensity analysis. *Educ. and Psych. Meas.*, Vol. 7, pp. 247-280.
- HYDE, K., et al. (1956). *Sunday Schools Today*. Free Church Federal Council.
- LODWICK, A. R. (1958). *An Investigation of the Question whether the Inferences that Children Draw in Learning History correspond to the Stages of Mental Development that Piaget postulates*. University of Birmingham Dep. Ed. dissertation (unpublished).
- PEEL, E. A. (1959). Experimental examination of some of Piaget's Schemata Concerning Children's Perception and Thinking, and a discussion of their educational significance. *Brit. J. Ed. Psychol.*, Vol. XXIX, Part II, pp. 89-103.
- PEEL, E. A. (1961). *The Growth of Pupils' Judgments—Thinking Comprehension*. Unpublished account of research at Princeton, New Jersey.
- STOUFFER, S. A., and GUTTMAN, L. (1950). *Measurement and Prediction in Studies of Social Psychology in World War II*. Vol. 1. Princeton University Press.

IX.—APPENDIX.

STORY A: MOSES AND THE BURNING BUSH (*Exodus* iii, 1-6).

A man called Moses was one day looking after a flock of sheep in a rather lonely place, close to a mountain. Suddenly, an angel appeared to Moses in a flame of fire, out of the middle of a bush. The curious thing was that the fire was burning away, but the bush itself wasn't burnt.

Moses said to himself, "I must go and look at it closer, to see why the bush isn't burned." Now when God saw Moses come nearer to the bush, God called out from the middle of the bush, "Moses! Moses!" and Moses, not knowing who it was calling, said, "Here I am."

And God said, "Come no closer and take off your shoes. You are standing on holy ground." Then God spoke again and said, "I am your father's God, and the God of great men like Abraham and Isaac and Jacob."

Then Moses hid his face, for he was afraid to look at God.

Question 1: "Why was Moses afraid to look at God?"

Pre-operational answer: "He hadn't spoken politely to God."

Concrete operational answer: "Moses thought God would chase him out of the holy ground because he hadn't taken off his shoes."

Formal operational answer: "The awesomeness and almightiness of God would make Moses feel like a worm by comparison."

Question 2: "Why was the ground on which Moses stood holy?"

Pre-operational answer: "Because there was grass on it."

Concrete operational answer: "It was where God was standing."

Q: How did that make it holy? "Because the holy would go down through God's feet into the ground and make it holy."

Formal operational answer: "The presence of God would hallow it like a magnetic field. The magnetic field is everywhere but the pole is in one spot. God is concentrated there."

Question 3: "How would you explain the bush burning but not being burnt?"

Pre-operational answer: "Some bad men and boys came and lit it again. The flowers, leaves and branches weren't burnt because they are up high."

Concrete operational answer: "God was protecting it with some watery spirit stuff."

Formal operational answer: "Instead of God appearing in person, he came in a bush. It seemed to be to his eyes but it wasn't really." *Q:* How do you mean? "He could have imagined it in his mind."

STORY B: CROSSING THE RED SEA (Exodus xiv).

Once, long ago, there lived in Egypt a people who were called Israelites, and they were made to work as slaves by the Egyptians. The Israelites were treated very cruelly, until their leader—Moses—persuaded the Egyptian king to let the slaves go free.

Then Moses led the Israelites out of Egypt, across the desert for many miles, until at last they camped on the shore of the Red Sea. Meanwhile, the king of Egypt had changed his mind, and was very angry that he had let the slaves go free. So he came after them with his army of six hundred chariots. Now, when the Israelites saw the army coming after them, they were afraid, but Moses said that God would save them.

Then God told Moses to stretch his hand over the sea. And at that very moment the waters parted, and the Israelites went across the sea on dry land to the other side.

When they were safely across, the Egyptian chariots started to come after them, but God told Moses to stretch his hand over the sea again. And at that moment the waters came together and the entire Egyptian army was drowned. And when the Israelites saw that they were saved, they feared God and believed in Him and in Moses, His servant.

Question 4: "How would you explain the waters of the sea dividing?"

Pre-operational answer: "The man ran past the blue sea and the white sea."

Concrete operational answer: "The blue sea went on one side and the white sea the other side." "God's palms (of his hands) were pushing them aside. You couldn't see them because they were invisible. When the Israelites were through he took his hands away and the waters flooded back."

Formal operational answer: "It's been proved at a special time the sea does part. It gets very shallow." *Q:* How do you mean? "At a special time of the year, once a year." *Q:* Did God do it? "No, they just got there at the right time. God had nothing to do with the sea, but he knew about it and got them there at the right time."

STORY C : THE TEMPTATIONS OF JESUS (*Matthew iv, 1-11, Luke iv, 1-13*).

When he was thirty years old, Jesus was led by the Spirit into the desert to spend forty days, where he was tempted by the devil. He ate nothing during that time and felt very faint and hungry.

Then the devil came to him and said, "If you really are the son of God, tell this stone to turn into a loaf of bread." Jesus answered, "The Scriptures say, Man shall not live by bread alone."

Then the devil took him up and showed him all the kingdoms of the world, and said to Jesus, "I will give you all this power and wealth that you see, if you will only fall down and worship me." Jesus replied, "It is written in Scripture, You shall worship the Lord your God, and Him only shall you serve."

Finally, the devil took him to Jerusalem, right to the top of the highest tower of the temple, and said, "If you really are the son of God, throw yourself down from here, for it says in Scripture that angels will take care of you." To this Jesus replied, "It is also written in Scripture, You shall not tempt the Lord your God."

And when he had tried every kind of temptation on Jesus, the devil went away for a time.

Question 5 : "Why wouldn't Jesus turn the stone into bread?"

Pre-operational answer : "Because he said not to eat bread alone."

Q : How do you mean? "They should have something else like cheese and something to drink."

Concrete Operational answer : "Because Jesus didn't want to show off."

Formal operational answer : "Jesus lives by the word of God."

Q : How do you mean? "You have faith in him and he will provide all your needs."

(*Manuscript received October, 1964.*)

TEACHERS' AND PARENTS' CONCEPTION OF THE TEACHER'S ROLE

By

F. MUSGROVE

(Department of Education, University of Bradford)

AND

PHILIP H. TAYLOR

(School of Education, University of Leicester)

SUMMARY. An inquiry was carried out among 470 teachers in grammar, modern, junior and infant schools to establish how widely or narrowly they conceived their role and what elements they gave greatest weight to within it. Grammar school teachers had a more constricted view of their role than their modern school colleagues, and junior school teachers working in middle-class areas than their colleagues in working class districts. Married women junior school teachers had a more restricted view of their role than single teachers.

In all types of school teachers saw their work primarily in intellectual and moral terms and were comparatively indifferent to more general social training, although modern school teachers gave more emphasis to social training than grammar school teachers. All saw parents as comparatively indifferent to moral objectives in the education of their children, primarily concerned with 'instruction' but also attaching great importance to 'social advancement.' A sample of 237 parents, in fact, gave a list of educational priorities which agreed in substance with the ratings made by teachers.

I.—INTRODUCTION.

CHANGES which are taking place or thought to be imminent in the contemporary teacher's role have been discussed in rather general terms by sociologists. Actual investigations of the teacher's role, the professional behaviour in which he engages and which is expected of him, have been few. Taylor (1962) reported an inquiry into pupils' expectations of teachers in different types of school; Musgrove (1961) reported an inquiry into parents' expectations of teachers in different types of urban area and at different social levels. The teacher works within a framework of expectations. He may respond to some of these expectations and reject others; he may misjudge the pressures brought to bear upon him and defend himself against demands which have not, in fact, been made. This paper reports an empirical inquiry into the way he sees his role in relation to the perceived and the actual expectations of parents.

Wilson (1962) has argued that the teacher's role must become more 'diffuse' at a time when most professional roles are becoming more specialized and specific. "The diffuse role means diffuse involvement." Mays (1962), has similarly argued that a teacher's role must broaden in scope, embracing ever more 'parental' functions and calling for the skills and interests of the social worker. "The argument then is between the ideas of the teacher as a pure inculcator of knowledge and the teacher as a welfare worker. It is not so much that the two interpretations of the teacher's role need be mutually exclusive. The disagreement concerns the degree of emphasis and the amount of time and energy to be devoted to these related aspects of the job."

II.—PLAN OF INVESTIGATION.

The purpose of the inquiry reported below was to estimate to what extent teachers in different types of school and circumstances saw their roles as 'diffuse' or 'restricted'; what weight they attached to different aspects of their work; what weight they thought parents expected them to give to different educational objectives; and what weight parents did, in fact, attach to these objectives.

The instrument used was a questionnaire which listed six commonly accepted educational aims. These aims which teachers are seen as pursuing were those most frequently mentioned by parents in the author's earlier inquiry (Musgrove, 1961). The aims were: I.—Moral Training (the inculcation of values and attitudes, e.g., honesty, kindness, tolerance, courage); II.—Instruction in Subjects (imparting information and promoting understanding of a body of knowledge); III.—Social Training (encouraging politeness, good manners, decency in speech and dress, etc.); IV.—Education for Family Life (training in human relationships with special reference to attitudes to the opposite sex); V.—Social Advancement (preparing children to 'get on in life'); and VI.—Education for Citizenship (developing an understanding of the modern world, etc.).

The teachers involved in this research were asked to rank these objectives as they valued them and also as their experience led them to believe that parents in general, valued them. They were also asked to indicate any objective which they regarded as no concern of theirs and none of their business as teachers. The parents who took part in the inquiry—a sample of 108 parents of primary school children and 129 parents of children at the secondary stage in one of the local authority areas selected for research—were asked to rank these objectives as they thought they should weigh with the teachers in charge of their children, and to indicate any of the listed objectives which they regarded as none of the teacher's business.

Grammar, modern, junior and infant schools were approached in three randomly selected local education authority areas. Four grammar schools co-operated (two mixed, one boys' and one girls'), and fourteen secondary modern schools (nine mixed, three boys' and two girls'). The schedule was completed by fifty men teaching in the grammar schools and by thirty-seven women; and by 103 men and ninety-one women in the modern schools.

Ten infant schools co-operated in the inquiry and seventeen junior (mixed) schools. One-hundred-and-eighty-nine teachers completed the questionnaire, fifty-one (all women) in infant schools, and 138 (forty-eight men and ninety women) in junior schools.

The primary schools were classified as denominational and non-denominational; and also according to the type of social area they served. (Scattered catchment areas made this classification impossible or unrealistic in the case of secondary schools.) Areas were classified as 'middle class' if some two-thirds of the parents were in white collar occupations, as 'working class' if two-thirds were manual workers; as 'mixed' if the proportions of white-collar and manual workers were approximately equal.

III.—RESULTS.

(1) *Diffuseness of role conception and expectation.*

An individual teacher's diffuse or restricted conception of his role was measured by the number of educational objectives he regarded as none of his

business. On this estimation grammar school teachers had a more restricted notion of their role than secondary modern school teachers: 9 per cent. of the ratings of male grammar school teachers indicated 'no concern of mine,' only 3.2 per cent. of the ratings by men teaching in modern schools ($C.R.=3.7$, $P<0.01$). The corresponding ratings by women in the two types of school were 11 per cent. and 4 per cent. ($C.R.=3.6$, $P<0.01$). The objectives which were rejected were exclusively social, as opposed to moral and intellectual aims.

Primary school teachers stood between grammar and modern school teachers in their restricted view of their role: 8 per cent. of their ratings indicated 'no concern.' Like teachers at the secondary stage, the aims they rejected were all social in nature.

TABLE 1

RESTRICTED OR DIFFUSE ROLE CONCEPTION OF JUNIOR AND INFANT SCHOOL TEACHERS
ACCORDING TO SEX AND MARITAL STATUS.

Teachers' Sex and Marital Status	Junior Teachers				Infant Teachers			
	N	Max. ranks	'No concern'	%	N	Max. ranks	'No concern'	%
Men	51	306	19	6.2				
Single Women	51	306	23	7.5	25	150	12	8.0
Married Women	35	210	46	12.4	22	132	11	8.3
All Women	86	516	49	9.5				
Total	138	328	68	8.2	47	282	23	8.2

TABLE 2

TEACHERS' RESTRICTED OR DIFFUSE ROLE CONCEPTION ACCORDING TO SOCIAL-CLASS AREA AND DENOMINATIONAL STATUS OF SCHOOL.

Type of Area and Denominational Status of School	Junior Teachers				Infant Teachers			
	N	Max. Ranks	'No concern'	%	N	Max. ranks	'No concern'	%
Middle Class and Mixed	66	396	42	10.6	32	192	16	8.3
Working Class	72	432	26	6.1	15	90	7	7.8
Denominational	40	240	23	9.6	14	84	6	7.7
Non-denominational ..	98	588	45	7.6	33	198	17	8.6

Infant school teachers, whether married or single, old or young, with or without much teaching experience, and irrespective of whether they worked in schools serving predominantly upper or lower-class social areas, rejected an und the same percentage of all objectives.

Married women teachers in junior schools rejected a significantly higher percentage of objectives than men teachers, single women teachers, and teachers in infant schools (C.R.=2.47, $P<0.02$, 1.99, $P<0.05$ and 2.35, $P<0.02$, respectively). (See Table 1.) Junior school teachers working in schools serving a predominantly middle-class catchment area have a significantly more restricted role conception than their colleagues in schools serving working-class districts. They rejected over 10 per cent., as against 6 per cent. of the objectives (C.R.=2.3, $P<0.02$). Married women junior school teachers and junior school teachers working in middle-class areas thus exhibited about the same degree of role restriction as teachers in grammar schools.

The 108 parents of primary school children who completed the questionnaire ascribed a significantly wider range of educational objectives to teachers if they were working class than if they were middle class. The opposite was the case with the 129 parents with children in secondary schools: working class parents ascribed to teachers a comparatively restricted role, particularly with regard to social objectives; middle class parents now expected a significantly wider range of services from teachers.

TABLE 3
ASCRPTION OF DIFFUSE OR RESTRICTED ROLE TO TEACHERS BY PARENTS.

	Percentage of ratings indicating objectives of no concern to teachers			
	'Social Class' (Registrar-General)			
	I-II	III Non-manual	III Manual	IV-V
Parents of Primary Children N : 108	11.8	9.7	6.3	2.8
Parents of Secondary Children. N : 129	1.6	5.2	10.6	9.2

(2) *Teachers' role conception, perceived expectations of parents, and actual expectations of parents.*

Teachers in all types of school saw their work primarily in intellectual and moral terms, placing greatest weight on instruction in subjects and moral training. They placed comparatively little emphasis on social objectives in general, and least of all on 'social advancement' in particular. In no type of school were teachers prepared to see themselves primarily as agents of social mobility. They saw parents as being comparatively indifferent to moral and social training, but placing great weight on instruction and on social advancement. In fact, the parents in general emphasised the same objectives as teachers: moral training and instruction in subjects; and, like teachers, gave comparatively little weight to 'social advancement' (although there were significant social-class differences in this regard).

There were no significant differences between the emphases of junior and infant teachers. Both gave greatest weight to moral training and least to social advancement; instruction in subjects and social training were given roughly equal weight after moral training.

Among teachers at the secondary stage, as at the primary, moral training took pride of place over 'instruction' for all except men teaching in grammar schools, who seem to perceive their job primarily in intellectual terms. As with primary school teachers, 'social advancement' came last except for women teachers in modern schools who, rather curiously, gave least emphasis to 'education for family life.'

TABLE 4
MEDIAN RANK ORDER OF TEACHERS' AIMS.
(Order ascribed to parents by teachers in brackets.)

	Moral Training	Instruction	Social Training	Family Life	Social Advance	Citizen-ship
Grammar School Men	2 (3)	1 (1)	4 (4)	5 (6)	6 (2)	3 (5)
Grammar School Women	1 (5)	2 (1)	5 (3)	4 (4)	6 (2)	3 (6)
Modern School Men	(1 (3)	2 (1)	3 (4)	5 (6)	6 (2)	4 (5)
Modern School Women . .	1 (3)	1 (1)	3 (4)	6 (5)	5 (2)	4 (5)
Actual parents' ranks	1	2	4	6	5	3
Junior Teachers	1 (4)	2 (1)	3 (3)	5 (6)	5 (2)	4 (5)
Infant Teachers	1 (3)	2 (1)	2 (4)	4 (6)	5 (2)	3 (5)
Actual parents' ranks	2	1	4	6	5	3

Some statistically significant differences emerged between the order of priorities found among modern school and grammar school teachers. Forty-two per cent. of men teaching in modern schools ranked 'social training' first or second, only 20 per cent. of men teaching in grammar schools did so ($X^2=7.03$, $P<0.01$). Similarly, with regard to moral training: 74 per cent. of modern school teachers gave it a high rating, only 60 per cent. of male grammar school teachers did so ($X^2=12.88$, $P<0.001$). Eighty-six per cent. of male grammar school teachers rated 'instruction' first or second compared with 63 per cent. in the modern schools.

Male grammar school teachers of non-academic subjects—woodwork, metalwork, religious instruction, art and music—gave no less weight to 'instruction' as their primary goal than teachers of academic subjects. Women teachers in the two types of school differed significantly in only one respect—the greater emphasis which secondary modern school teachers placed on social training as an educational aim. Only 5.6 per cent. of women teaching in grammar schools ranked this objective high, but a third of the women teachers in modern schools did so ($X^2=9.31$, $P<0.01$).

The following chart shows the educational priorities of all the teachers at the secondary stage compared with the priorities they ascribed to parents and with the priorities which were, in fact, found among the sample of parents of secondary school children. (There were 281 teachers and 129 parents.)

Per cent.
ranking
aims
high
(1 and 2)

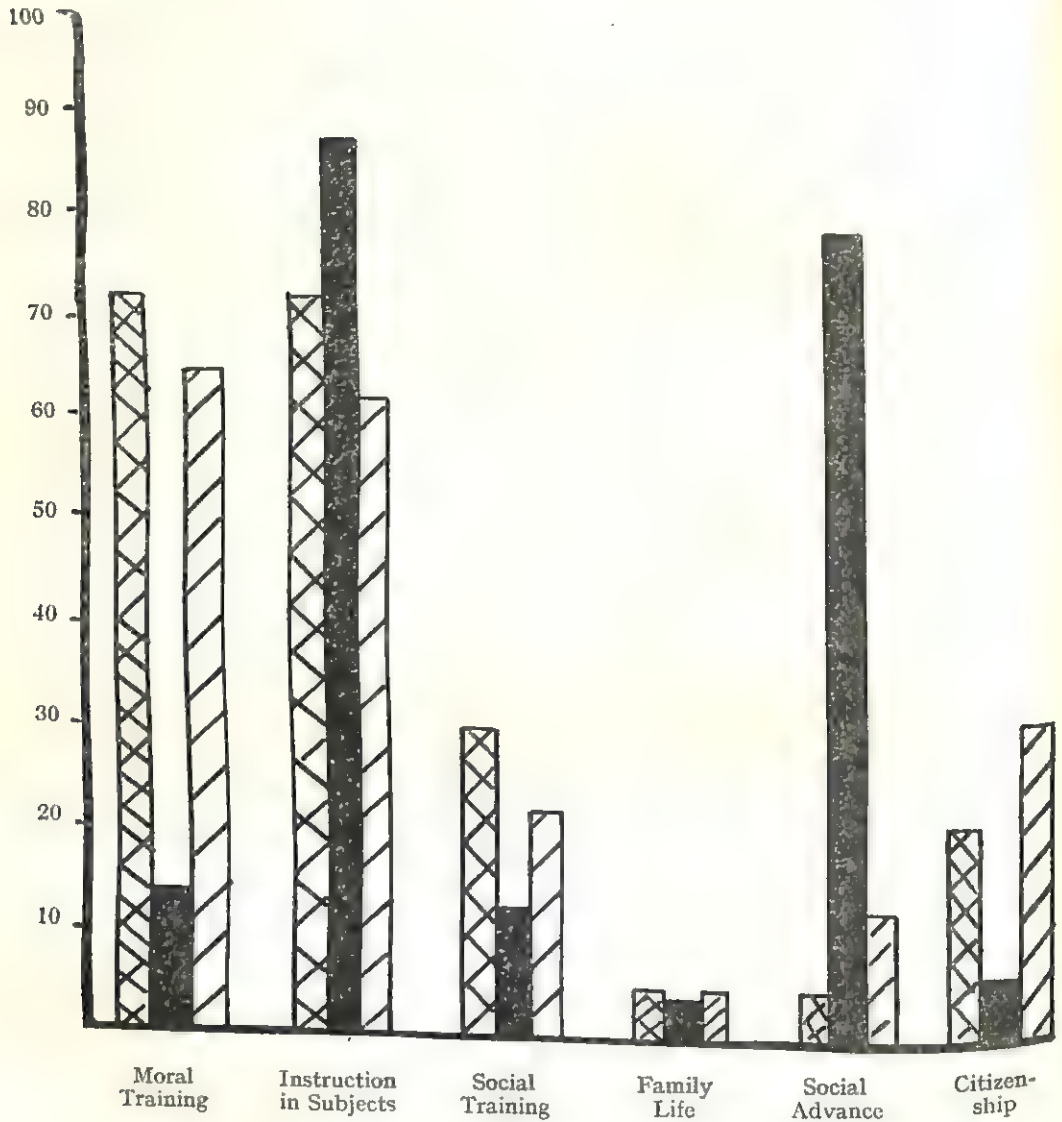


FIG. 1

TEACHER-EVALUATION, PERCEIVED PARENT-EVALUATION, AND ACTUAL PARENT-EVALUATION



TEACHER-EMPHASIS

PERCEIVED PARENT-EMPHASIS

PARENT-EMPHASIS

Parents, like teachers, in fact, gave greatest weight to moral training and instruction and comparatively little to social advancement. Working class parents, however, gave significantly more weight to social advancement than did middle class parents: 18 per cent. of the former ranked it high, only 2 per cent. of the latter ($C.R.=3.1$, $P<0.01$). There were no other significant differences between parents of different occupational standing. There was a trend for middle class parents to rank 'instruction' high (72 per cent. did so compared with 59 per cent. of working class parents) but this difference falls short of statistical significance.

IV.—DISCUSSION.

According to the measures used in this inquiry, teachers in grammar schools appear to have a more restricted notion of their role than teachers in modern schools: they see their work in moral and intellectual terms and to a marked degree reject social training as any part of their business. The more diffuse role conception of modern school teachers is not surprising in view of the short history and more uncertain and perhaps wider objectives of modern schools, and the grammar school's traditional concern with intellect and 'character.'

Junior school teachers working in predominantly middle class areas have a more restricted view of their role than teachers in working class districts. This probably reflects a realistic response to the social environment in which they find themselves, an acceptance of the need to provide social training which in socially better districts can safely be left to the home.

Married women teachers in junior schools seem less prepared to accept a diffuse role than their single colleagues, whatever kind of social district they serve. This was not the case with married women teachers in infant schools; and we might speculate that infant teaching in any case attracts women who are more prepared to accept a general 'mothering' function as a necessary aspect of their work.

Parents' 'diffuse' or 'restricted' expectations of teachers appear to be related both to their social class and to their children's stage of education. In a previous inquiry Musgrove (1961) found that working class parents of primary school children wished to place far greater responsibility for behaviour-training on teachers than middle class parents, who felt that this was the responsibility of the home. The responses of parents of primary school children in the present inquiry are in line with these earlier findings. At the secondary school stage it seems to be middle-class parents who would make the widest and most general demands on teachers (except that working class parents have stronger expectations with regard to 'social advancement'). Working class parents perhaps look particularly to the school as an ally when their children are young, vulnerable and exposed to an inauspicious local social environment.

Teachers in all types of school see their role in moral and intellectual terms and are comparatively indifferent to the more specifically social aims of education. Secondary modern school teachers placed more weight on social objectives than their grammar school colleagues, but nevertheless, placed much greater emphasis on 'instruction in school subjects.' In emphasising 'instruction' teachers were in line with what they thought parents expected and with what parents in fact expected.

But the area of discrepancy between teachers' aims and what they imagine to be parents' is still very large. On the whole, teachers take an unflattering view of parents (and their own aims are remarkably idealistic), seeing them as indifferent to moral training but very concerned with social advancement. In fact, parents were substantially in agreement with teachers. The area of (unnecessary) tension might be considerably reduced if parents and teachers established more effective means of communication.

V.—REFERENCES.

- MAYS, J. B. (1962). *Education and the Urban Child*. Liverpool: Liverpool University Press.
- MUSGROVE, F. (1961). Parents' expectations of the junior school. *Soc. Rev.*, 9, 167-180.
- TAYLOR, P. H. (1962). Children's evaluations of the characteristics of the good teacher. *Brit. J. Educ. Psychol.*, 32, 258-266.
- WILSON, B. R. (1962). The teacher's role—a sociological analysis. *Brit. J. Soc.*, 13, 15-32.

(Manuscript received 7th September, 1964.)

PSYCHOLOGICAL ASPECTS OF CRAFTSMANSHIP IN POTTERY-MAKING AT A SECONDARY SCHOOL

By G. H. JAMIESON*

(*Department of Education, University of Manchester*)

SUMMARY. A course of training in pottery making was given to 223 pupils in the third forms of a comprehensive school. At the end of the course, the pupils were tested in their ability to make pottery in accordance with a standard model. In addition, a battery of tests was administered comprising verbal, non-verbal, spatial, interest, personality and tactile-kinaesthetic tests.

The results indicated that the N.I.I.P. Spatial Test 81 had the highest predictive value. It was also found that intercorrelations between the verbal, non-verbal, spatial, tactile-kinaesthetic and pottery criterion tests were all significantly positive.

It was further found that neuroticism, as measured by the M.J.P.I., was negatively (and significantly) correlated with the criterion of skill. No relationship was found between practical interests and the criterion, but there was a significant positive correlation between extraversion and practical interests.

I.—INTRODUCTION.

FACTORS affecting the development of practical skills in school children are largely unknown; apart from the work of Fitzpatrick (1951), this has been a neglected field of study. Educational research bearing any relationship to the problem has been concerned mainly with whether practical ability is (Alexander, 1935), or is not (Price, 1940), an independent factor.

Prior to the second world war, research was directed to the motor aspect of skills, but post-war developments (Bartlett, 1951; Annett and Kay, 1956; Welford, 1958) give greater emphasis to its perceptual component. A further development has been the attempt to measure perceptual 'load' (Crossman, 1956) in terms of information theory. The technique evolved by Crossman would appear to offer interesting possibilities for further research into more complex sensori-motor skills.

The effect of personality on learning skills is also largely unknown, although it has been shown by Lynn and Gordon (1961) that introverts are slow and accurate at manual tasks, whereas extraverts are quick and inaccurate.

The aims of the present research were to assess new empirical evidence on the ability of children to learn the skill of pottery-making, and to relate the findings to scores obtained from verbal, non-verbal, spatial, interest, personality, and tactile-kinaesthetic tests.

II.—METHOD.

(a) *Training.*

The pupils were given individual tuition in the use of powered and manual potter's wheels. Each pupil followed a set course of training, comprising twenty 1½ hours weekly practice lessons. The pupils were restricted to 12-oz. balls of clay during practice.

* Now of M.R.C. Unit, Department of Psychology, University of Liverpool.

(b) *Task.*

At the end of the training period, each pupil was given the task of making pottery to conform in size and shape to a specimen pot which had a fixed diameter of 3-in., and a height of $2\frac{1}{2}$ -in. A uniform 12-oz. ball of clay was given to each pupil for the test. The finished work was measured with callipers in discrete steps of $\frac{1}{8}$ -in., deviations from the standard model were recorded as negative values.

(c) *Tests.*

- (1) N.I.I.P. Group Verbal Test 36.
- (2) N.I.I.P. Group Non-Verbal Test 75.
- (3) N.I.I.P. Group Spatial Test 81.
- (4) Maudsley Junior Personality Inventory.
- (5) Devon Interest Test.
- (6) Tactile-kinaesthetic tests :
 - (i) Discrimination of firmness.
 - (ii) Discrimination of shape.
 - (iii) Discrimination of thickness.

(d) *Subjects.*

223 pupils (119 boys and 104 girls) from eight third-year forms of a comprehensive school in a large industrial city.

III.—RESULTS.

TABLE 1

CORRELATION MATRIX FOR NINE VARIABLES.

	1	2	3	4	5	6	7	8	9
1	—	53	43	15	-16	-16	01	27	36
2	53	—	55	-07	-02	-18	00	27	33
3	43	55	—	00	19	-21	-06	23	45
4	15	-07	00	—	-43	00	-13	05	01
5	-16	-02	19	-43	—	-07	22	-05	08
6	-16	-18	-21	00	-07	—	02	-06	-22
7	01	00	-06	-13	22	02	—	07	-02
8	27	27	23	05	-05	-06	07	—	28
9	36	33	45	01	08	-22	-02	28	—

Key to variables :

1. N.I.I.P. Verbal Test 36.
2. N.I.I.P. Non Verbal Test 75.
3. N.I.I.P. Spatial Test 81.
4. Devon Interest Test 'A' Score.
5. Devon Interest Test 'P' Score.
6. Maudsley J.P.I. 'N' Score.
7. Maudsley J.P.I. 'E' Score.
8. Tactile-kinaesthetic Tests.
9. Pottery Criterion Test.

Above 14 significant at the 5 per cent. level.

Above 18 significant at the 1 per cent. level.

A correlation matrix was compiled, see Table 1, from the results of nine variables. Five of these variables gave significant correlations with the criterion; of these, the highest correlations were with the cognitive tests. The highest correlation with the pottery criterion test was that given by the spatial test. Neuroticism gave a correlation with the criterion which was negatively significant (1 per cent. level). Moreover, it was also found to have significant correlations in the same direction, with the cognitive tests. The Interest Test revealed no significant correlation between practical interests and the criterion.

TABLE 2
CENTROID FACTOR MATRIX OF NINE VARIABLES.

Variable	Factor					h ²
	1	2	3	4	5	
1	52	47	-21	-04	04	54
2	65	30	-22	29	07	65
3	68	25	24	26	-19	68
4	-23	55	05	-23	11	42
5	31	-59	26	12	-17	55
6	-29	-08	-18	06	-17	15
7	15	-31	-20	-20	20	24
8	36	19	-14	-14	10	21
9	57	23	21	-17	04	45

NOTE.—For key to variables, see Table 1.

A factor analysis by the centroid method was carried out, and the results are given in Table 2. It will be seen that academic interest and neuroticism both show negative loadings on the first centroid factor, whereas, on the second centroid factor, practical interest and extraversion have negative loadings.

TABLE 3
ANALYSIS OF VARIANCE BETWEEN FORMS.

Test	'F' ratio	Significance level
Verbal	24.6	1%
Non-verbal	11.1	1%
Spatial	13.4	1%
Academic Interest	1.2	N.S.
Practical Interest	1.8	N.S.
Neuroticism	3.1	N.S.
Extraversion	2.0	N.S.
Tactile-kinaesthetic	4	1%
Pottery	20.4	1%

NOTE.—Significant results all favoured the higher forms.

Analyses of variance were carried out on the data from each test, and the 'F' ratios and significances are shown in Table 3. In all cases where the results are significant, the differences are in favour of the higher forms.

IV.—DISCUSSION OF RESULTS.

The aim of this investigation was to find out by means of controlled experiments the main psychological factors which play a part in the performance of skill in pottery-making, with particular reference to the abilities of children in a comprehensive school.

It was found that the cognitive tests gave the highest correlations with the criterion. This suggests that the children with the best reasoning ability were also likely to be above average at pottery-making. An explanation, in terms of recent psychological theory on skills, could be given that children with superior intellect are more capable of interpreting perceptual cues, but further research is clearly necessary to establish the strength of this relationship.

It was hoped that the tactile-kinaesthetic tests would throw fresh light on the relationship between skill and sensory discrimination; but from the results obtained, it would appear that there is little to be gained from administering time consuming tests of this nature for purposes of prediction. This finding confirms the suggestion made by Vernon and Parry (1949) that more successful selection for special aptitudes can be achieved with the aid of paper and pencil tests.

The results show that the spatial test yielded the highest correlation with the criterion. This finding is in line with other investigations in this field (summarised by Vernon and Parry, 1949) which show good validities between spatial-mechanical tests and practical work.

Possibly the most interesting finding from this research was the negatively significant correlation between neuroticism and the criterion. Others (Bartlett, 1947; Annett and Kay, 1956; Seymour, 1962) have stressed the importance of stability in the acquisition of skills, but evidence is offered here that emotional stability may also affect ability to learn practical skills.

The results also show that a child's expressed interest in practical work cannot be used as a reliable criterion for predicting successful performance in a craft skill.

V.—REFERENCES.

- ALEXANDER, W. P. (1935). Intelligence, concrete and abstract. *Brit. J. Psychol.*, Monogr. Suppl., No. 19.
- ANNETT, J., and KAY, H. (1956). Skilled performance. *Occup. Psychol.*, 30, 112-117.
- CROSSMAN, E. R. F. W. (1956). *The Measurement of Perceptual Load in Manual Operations*. Unpublished Ph.D. Thesis, University of Birmingham.
- BARTLETT, F. C. (1951). The bearing of experimental psychology upon human skilled performance. *Brit. J. Industr. Med.*, 8, 209.
- FITZPATRICK, T. F. (1951). *The Construction of a Test of Practical Ability with Special Reference to Woodwork*. Unpublished M.Ed. Thesis, University of Manchester.
- LYNN, R., and GORDON, I. E. (1961). The relation of neuroticism and extraversion to intelligence and educational attainment. *Brit. J. Educ. Psychol.*, 21, 194-203.
- PRICE, E. J. J. (1940). The nature of the practical factor (F). *Brit. J. Psychol.*, 30, 341-351.
- SEYMOUR, W. D. (1962). Experiments on the acquisition of industrial skills. *Occup. Psychol.*, 36, 15-21.
- VERNON, P. E. (1949). The structure of practical abilities. *Occup. Psychol.*, 23, 81-96.
- VERNON, P. E., and PARRY, J. B. (1949). *Personnel Selection in the British Forces*. London: Univ. London Press.
- WELFORD, A. T. (1958). *Ageing and Human Skill*. London: Oxford Univ. Press.

(Manuscript received 16th November, 1964.)

CLASS INFLUENCES IN HIGHER EDUCATION

By C. T. SANDFORD, M. E. COUPER AND S. GRIFFIN

SUMMARY. A survey by questionnaire, completed by 414 students at Bristol College of Science and Technology, and designed to provide a profile of student background and activity, was followed by interviews with ninety-seven of these students. The interviews were concerned primarily with student motivation to higher education and especially with the influence of home background. Analysis revealed that there was a substantial proportion of students, largely from working class homes, whose parents were indifferent or hostile to their children's higher education, and for whom the C.A.T. formed the highest acceptable aspiration; that working class students who lacked family support for higher education and who had not been motivated to continue their education by school or peer groups, had formed this aspiration after experience in the work situation and had often received the necessary support from their firms; that some working class students found higher education in a sandwich course at a C.A.T. acceptable because of its job orientation, whereas a full-time university course was not; and that lack of confidence and lack of knowledge were also factors inhibiting students whose families had no experience of higher education from applying to a university.

I.—INTRODUCTION.

THE difficulties of young people from working class homes in going to university are well known. Studies have shown these difficulties and also pointed to the wastage alluded to in *Early Leaving* (H.M.S.O., 1954) of those who were unable to continue their education. (Floud, Halsey and Martin, 1957; Crowther Report, 1959; Jackson and Marsden, 1962). The research undertaken at the Bristol College of Science and Technology was seeking to find if the degree level sandwich course at a College of Advanced Technology was a form of higher education which such students could more readily undertake—if it was a form of higher education acceptable to students who were unable to aspire to a university education, and more acceptable than university to students who would have entered upon a course of study at university with misgivings.

The research was also concerned to test the hypothesis that the family background of students at this C.A.T. had been the most important influence in motivating them towards higher education or in making them initially doubt its value (Kahl, 1953).

II.—METHOD.

In March, 1963, a survey was undertaken of all students at the Bristol College of Science and Technology who were studying for qualifications equivalent to first degrees. The survey, by questionnaire, covered 414 students, of whom 299 were reading for the Diploma in Technology—in Applied Biology, Applied Chemistry, Aeronautical, Electrical and Mechanical Engineering—sixty-nine were students of Pharmacy and forty-six of Architecture. The questionnaires were completed anonymously by the students under group supervision; the overall response rate was 95.2 per cent., and for no course was the response rate for the questionnaire as a whole less than 92 per cent.

The survey included questions on student educational background, choice of a C.A.T. education, place of residence when at College, financial circumstances, social activities, religion, politics, and parental occupation and education. The

survey was followed by interviews to investigate further the factors which influenced the students to seek higher education in general and a place at a C.A.T. in particular, and especially the part played by home environment. The students were selected for interview by systematic sampling within a population stratified by course and year. The intention was to interview 25 per cent. of the students who completed the questionnaire, but fourth-year Aeronautical, Mechanical and Electrical Engineering students left the College before the interviews could be undertaken, and it was not possible to interview three other students, so there was a loss of 11 per cent. of the sample; interviews, therefore, covered 23.7 per cent. of the total population of the survey.

TABLE 1
DISTRIBUTION OF STUDENTS IN SURVEY AND INTERVIEW SAMPLE.

Year	App- lied Biology	App- lied Chem- istry	Engineering			Total Dip. Tech.	ARIBA	Pharm- acy	Total
			Aero- nautical	Mech- anical	Elec- trical				
1. Survey	18	16	10	24	28	96	11	32	139
Interview Sample ...	5	4	3	6	7	25	3	9	37
2. Survey	14	15	11	16	29	85	17	22	124
Interview Sample	4	4	3	4	8	23	5	5	33
3. Survey	9	13	17	11	18	68	13	15	96
Interview Sample	2	3	3	2	5	15	3	4	22
4. Survey	7	7	20	7	9	50	5	—	55
Interview Sample	2	2	—	—	—	4	1	—	5
Total :									
Survey	48	51	58	58	84	299	46	69	414
Interview Sample	13	13	9	12	20	67	12	18	97

The main concern was to gain information about students reading for the Diploma in Technology as this is a new type of award and less is known about these students than university undergraduates. The study was less concerned with the Pharmacy and Architecture students whose courses differ from the Diploma in Technology in several ways which may affect recruitment to them. The Pharmacy course is full-time and three years' in length, and Architecture a five-year sandwich course entirely college based* with entry requirements only recently raised to the level of normal university entrance. All students in both these courses were studying for qualifications awarded by their appropriate professional bodies, except fourteen Pharmacy students studying for London B.Sc. degrees. These qualifications are of much longer standing than the Diploma in Technology and the courses leading to them are advertised in the literature of the professional associations.

Eleven of the students interviewed were female, of whom five were on Diploma in Technology courses; all but two students were single; of the Diploma in Technology students twenty-five (37 per cent.) were college based and forty-two (63 per cent.) were industry based.

* Students whose practical placements are arranged by the College. Industry based students return to their sponsoring firm for practical placements.

Social class background.

Classification of the students in the survey by father's occupation using the Registrar-General's five grade classification revealed a relatively larger proportion of students from working class homes at the C.A.T. than normally found at provincial universities. Comparison is made in Table 2, with the background of students from Nottingham University, similarly classified. (Allen, *et al.*, 1962). Nottingham was selected for comparison, as it was a 'redbrick' university without the attractions of Oxbridge and London, and like the Bristol C.A.T. drawing students from the whole country but with a regional bias. The Bristol and Nottingham surveys were conducted within 12 months of each other. It will be seen that a much larger proportion of students at the C.A.T.

TABLE 2
SOCIAL CLASS ANALYSIS BY OCCUPATIONS OF FATHERS.
(Registrar-General's Classification.)

Social Class	Fathers of BCST students, 1963		Fathers of Nottingham University students 1962*	Percentage Occupied and Retired Males in England and Wales, 1951
	Total	Diploma in Technology		
I	9.2	10.0	18.1	3
II	30.9	26.9	46.4	14
III	46.0	49.5	31.7	52
IV	6.8	6.7	} 2.0	16
V	2.4	2.7		15
Don't know	4.6	4.0		—
Total	100	100	100	100

*—Although stated to be Registrar-General's Classification 1-5, the Nottingham Survey in fact used a regrouping of these classes for which correction is made in these figures.

came from social class III and smaller proportions from social classes I and II than at the university: although the C.A.T. still had over twice the proportion found in the population at large in the latter classes, the proportion from class III was close to that for the nation as a whole. Classes IV and V were under represented at both institutions, but the University drew less students from these classes than did the C.A.T. The value of such analysis is, however, partly vitiated by the wide occupational range of the Registrar-General's classification especially in class III.

In the analysis of the interviews, the students were divided into two classes, working class and middle class. Students categorised as working class were those whose fathers' occupations came within the Registrar-General's classification of social classes V, IV, III (manual), and III (non manual) where either the father had not received a secondary education or where the father's first job on leaving school was classified as social class III (manual), IV or V. Students who were classified as middle class had fathers whose occupations fell in the reduced social class III (non-manual), and social classes II and I (see Table 3).

TABLE 3

ANALYSIS OF STUDENTS INTERVIEWED BY SOCIAL CLASS.

Class	Diploma in Technology	ARIBA and Pharmacy	Total
Middle Class	25 (32%)	16 (53%)	41
Working Class	41 (68%)	14 (47%)	55
Total	66 (100%)	30 (100%)	96*

* One student interviewed has been excluded from the analysis as he only came to this country at the age of 16 years.

This report concentrates mainly on the Diploma in Technology students interviewed, the students in the other two categories being included only for the purposes of comparison.

Home Environment and Higher Education.

In analysing the influences of home background, three categories were distinguished:

TABLE 4

PARENTAL ATTITUDES TO HIGHER EDUCATION OF STUDENT.

Category	No. of Students (Dip. Tech.)
1. Those who received active support and encouragement from their parents to continue their education	39
2. Those whose families, while not actively discouraging continued education, did not offer support	22
3. Those whose parents were against the continuation of their education and who were at college despite the influence of their family	5

Of the Diploma in Technology students interviewed, thirty-nine were in the first category, twenty-two in the second and five in the third. When these categories were related to social class (Table 5) a much larger proportion of working class students were in the last two categories (Chi-squared=3.699. Probability almost .05).

Relative importance of influences to Higher Education.

The students were questioned about the relative importance of the family, the school, peer groups and work in motivating them to higher education. As was expected (Kahl, 1963), for the majority the family played the main motivating role. For all those Diploma in Technology students whose wish for higher education was supported by their parents, the family had been the

main source of motivation. On the other hand, for twelve of the Diploma in Technology students whose parents had not supported higher education, the influence of the family had worked in the other direction; they formed no aspirations for higher education whilst at school, and only were motivated to follow such a course of study after experience in the work situation. Six of these students were eligible for university entry. Where the family had been the main source of motivation to higher education, there was no significant difference between the number of cases where the mother had been the main influence and those where the father's influence had been more important.

TABLE 5

HOME ENVIRONMENT AND SOCIAL CLASS OF DIPLOMA IN TECHNOLOGY STUDENTS.

Parents' Attitudes	Parents' Social Class		Total
	Middle Class	Working Class	
Support higher education (Category 1)	19	20	39
Do not support higher education	6	21*	27
Total	25	41	66

* Including parents of five students who actively opposed higher education.

For only two of the Diploma in Technology students was the school the main influence to continued education (although it had been the main influence for six of the students in Architecture and Pharmacy). Most students expressed the opinion that they had felt that the school had favoured pupils continuing to higher education, but that for them it had played a negligible role in motivation. There were cases where the school had allowed a boy to take a combination of subjects which disqualified him from university entry.

Neither in school nor at home had peer groups played a major role in persuading a student either to continue his education or to go out to work on leaving school. However, for many students the grammar school peer groups in particular had reinforced the opinion that higher education was important, and had played a useful supporting role, especially where the student was the first in his family to continue his education beyond school.

Aspiration to Higher Education whilst still at school.

The Diploma in Technology students were asked to recall their attitude to higher education at the age of about 15, the earliest age at which they could have left school and the age at which most of them were beginning to prepare for their G.C.E. 'O' level examinations. They were asked if at that age they had considered continuing their full-time education beyond school or were expecting to start work, either at 15 or 16 years or at 18 years, and also if they were aiming specifically at university by the age of 15. The results are summarised in Tables 6 and 7.

TABLE 6

ATTITUDE OF DIPLOMA IN TECHNOLOGY STUDENTS TO HIGHER EDUCATION AT THE AGE OF ABOUT 15 YEARS.

Student aspiration at 15 Years	Parents' Social Class		Total	Parents' attitudes to Higher Education		Total
	Middle Class	Working Class		Support	No Support	
Aimed at higher education	21	20	41	30	11	41
Did not aim at higher education	4	21	25	9	16	25
Total	25	41	66	39	27	66
Significance	$X^2=6.75, P<.01$			$X^2=7.19, P<.01$		

TABLE 7

ASPIRATION OF DIPLOMA IN TECHNOLOGY STUDENTS TO UNIVERSITY ENTRANCE AT THE AGE OF ABOUT 15 YEARS.

Student aspiration at 15 Years	Parents' Social Class		Total	Parents' attitudes to Higher Education		Total
	Middle Class	Working Class		Support	No Support	
Aimed at university	19	5	24	21	3	24
Did not aim at university	6	36	42	18	24	42
Total	25	41	66	39	27	66
Significance	$X^2=24.64, P<.001$			$X^2=12.397, P<.001$		

Of nineteen middle-class students with parents supporting their aspiration to higher education, at the age of 15 years, eighteen were aiming at higher education and seventeen specifically at university education. But of the twenty working-class students whose parents support their ambition for higher education, at the age of 15 years only twelve had considered higher education and of these only four were aiming at university entry. At a time when a large majority of the middle class students were not only aware of the possibility of going to university, but were actively working towards that end in beginning to prepare for their 'O' level examinations only about half the working-class students had considered that they might extend their education after leaving school, and of these very few had in mind the possibility of going to university. It became clear from the interviews that the middle-class students when at

school saw their school careers in a wider context of continuing education—to stay on at school and continue to higher education was normal, to leave unusual. For the working-class student to stay on was unusual, and during their school careers many of these students were aiming only at the next examination hurdle, first 'O' level, then 'A' level. On doing well at one stage, they began to consider the next, and sometimes found themselves ill-equipped, as subjects neglected at 'O' level left them ineligible for university after 'A' level. Many of the working class students appear to have arrived on their present courses having been passed from stage to stage of the education machine without much appreciation of what came next, whereas the middle-class students were aiming more deliberately at longer term goals.

University Eligibility and Application.

The survey by questionnaire revealed that 270 (65.5 per cent.) of the students said that they were eligible for university entry. Of these, 144 (57 per cent. of the 270 students eligible) had made no application to a university.

One-hundred-and-twenty-two students had applied to at least one university and, of these, thirty withdrew their applications on acceptance by the Bristol College of Science and Technology, forty-three were offered at least one place which they turned down and only forty-nine students had all their university applications rejected. There was, therefore, a large body of students who appeared to prefer the C.A.T., to a university although they were eligible to apply for a place there. This was investigated further in the interviews of Diploma in Technology students and the results classified according to social class and parents' attitudes. There was no significant difference between the proportions of working-class and middle-class students eligible for university entry, nor was student eligibility for university entrance significantly related to the support for higher education from their parents.

TABLE 8

ELIGIBILITY OF DIPLOMA IN TECHNOLOGY STUDENTS FOR UNIVERSITY ENTRY.

Students' eligibility for University	Parents' Social Class		Total	Parents' attitudes to Higher Education		Total
	Middle Class	Working Class		Support	No Support	
Eligible	13	22	35	21	14	35
Not Eligible	12	19	31	18	13	31
Total	25	41	66	39	27	66

There were differences, however, between class groups and groups with and without parental support who applied to university, which although not significant, are in the expected direction (see Table 9).

It is useful to look in more detail at the students in these categories. Considering first the Diploma in Technology students whose parents supported

their ambition for higher education, of the twenty working-class students in this category, nine were eligible to go to university. Two of these were, in fact, unable to secure university places. Of the remaining seven, five did not apply to a university and two applied but turned down places in order to come to the College. A large element in the behaviour of these seven appears to have been lack of confidence. One mentioned that his parents were suspicious of universities and the effect they might have on their son, and that the C.A.T. enabled him to receive higher education without the 'stigma' of going to a university. Also an emphasis was placed by these students on job training, which was considered by both the students and their families to be more important than being educated at a 'prestige' institution.

TABLE 9
APPLICATIONS TO UNIVERSITY OF DIPLOMA IN TECHNOLOGY STUDENTS.

Student University Applications	Parents' Social Class		Total	Parents' attitudes to Higher Education		Total
	Middle Class	Working Class		Support	No Support	
Applied	12	7	19	14	5	19
Did not apply	13	34	47	25	22	47
Total	25	41	66	39	27	66

Twelve of the nineteen middle-class students in this category were eligible for university entry, four of these being unsuccessful in their attempts to secure places. Four of the students who might normally have been expected to try for university places and who either did not apply or who withdrew their applications seemed to be reacting against strong parental direction. All of the remaining four students applied to universities but withdrew their applications on acceptance by this College through lack of confidence in their ability to gain or maintain a university place.

Of the Diploma in Technology students whose parents did not support their desire for higher education, sixteen came from working-class homes, including eight who were eligible for university places. Only one applied to a university and he withdrew his application on acceptance by the College, sharing his parents emphasised view on the importance of job training. The seven students who were eligible for but did not apply to universities gave many reasons, very often not too certainly, for applying only to the C.A.T. The reasons included lack of knowledge of possibilities, the feeling that a practical type of course better suited 'people like them,' finance, lack of confidence (no one in the family having done anything like it before) and rejection of a continuation of 'middle-class education' after secondary education. Only two of the six middle-class students in this category were eligible for university places. Both of these applied but withdrew their applications, one for financial reasons and one because she could not study the subject of her choice at university.

Three of the five Diploma in Technology students whose parents opposed their continued education and who have entered higher education, despite their parents' wishes were eligible to go to university. One of these students, despite being awarded a State scholarship, did not apply to a university as he doubted his ability to cope with a university course. The other two students applied to a university but withdrew their applications, emphasising the 'work' nature of their present courses in their comments. The parents of one of these students looks on his period at college as part of his job.

Four of these five students are aware of increasing social distance between them and their families, and of strained relationships. The one student still on good terms with his family, emphasised the 'work' image of the Diploma in Technology courses, sharing his parents' view rejecting universities and describing all at universities as 'wet behind the ears.'

For eighteen students, fourteen of them from working-class homes, and twelve of them with parents who did not support their higher education, the 'job training' bias of the course, that they were following was important. They considered their present courses worth while because they gave training for a specific job, and a paper qualification, which was necessary today, but they would not have been prepared to devote the time to a university course that provided a less job-orientated education.

Knowledge and confidence.

Only seven of the Diploma in Technology students interviewed had parents who had received any form of higher education, all of them middle-class students whose parents supported their wish for higher education. It became very clear that many students had felt unsure about embarking on a course of higher education, largely because no-one in the family had undergone such an experience before. Lack of confidence, evident in many students, was a major factor for sixteen of the sixty-six Diploma in Technology students interviewed, causing them not to apply to university, turn down university places or even not ensure their eligibility for university entrance. Nine of these students came from working-class homes, seven from middle-class homes, and they include eleven whose parents supported their ambition for higher education. Eleven of the sixteen were the eldest in their families.

Many students felt that they had been handicapped, both in making themselves eligible for university entry and in knowing what opportunities were open to them by the lack of knowledge of procedures and requirements, on their own part and that of their families, again because no-one from home had undertaken higher education before. This was particularly true of twenty-four students, eighteen of whom came from working-class homes. This accords with experience in America as well as in this country. In 1963, Ellis and Clayton wrote :

"Generally, lower class youth find themselves confronted by an environment in which going to college is the exception, not the rule, and in which strong counterpressures may be mounted against those who seek to deviate from the prevailing cultural norms. Consequently, if college goals are to emerge among lower class youth, some substitute channels must exist for transmitting information about college-information that is traditionally handed down to children in the more favoured classes."

Relationship with family.

Relationships with their families were discussed with each of the Diploma in Technology students and each student was classified on a three-point scale of a good relationship, a generally good relationship with some difficulties and anxieties, and a poor relationship. It was interesting to note that for most of the students in this last category, problems of communication and social distance had begun to arise before the student came to college, while he was at secondary school, but had increased during his period at college.

Bearing in mind Merton's hypothesis of the dissociative consequences of social mobility, it was particularly interesting to see if the students from working-class homes, now moving into middle-class occupations, had more strained relationships with their parents than middle-class students. The student parent relationships of those students whose parents were sympathetic to their higher education and those with apathetic parents were also compared. From these comparisons were omitted the five students whose parents were opposed to higher education, four of whom are now on bad terms with their families. Neither social class nor parental attitudes to higher education made any significant difference to the proportions of students with good and bad relationships with their parents (see Table 10).

TABLE 10
RELATIONSHIPS OF DIPLOMA IN TECHNOLOGY STUDENTS WITH PARENTS.

Student's Relationships	Parents' Social Class		Total	Parents' Attitude to Higher Education		Total
	Middle Class	Working Class		Support	No Support	
Good	14	19	33	23	10	33
Reasonably good but with anxieties	5	8	13	6	7	13
Poor	6	9	15	10	5	15
Total	25	36	61	39	22	61

Bearing in mind the few middle class students whose parents had had higher education, this step to C.A.T. is, of course, a big one for them as well as the working class student. Nevertheless, one might have expected the strain to have been greater for the working-class students and the unsupported students but, in fact, there is little difference in the proportions of bad relationships between either the class groups or the support groups. It might tentatively be suggested that the image of the course of study that the students are following in this type of college is one that is not so incompatible with their home environment as would be that of the university, and this does not put such a strain on their relationship with their parents. Indeed, one student said that his relationship with his family was better now that he had left a middle-class type of school.

For about a quarter of the students, whether analysed by class or degree of support, relationships with their families were poor. How far this was a consequence of their coming to college—or their previously having attended grammar schools—and how far a consequence of their growing up, irrespective of their education, is difficult to say.

IV.—CONCLUSIONS.

The survey and interviews reveal that despite significant class differences in attitudes to education (just over half the Diploma in Technology students had entered higher education without the active support of their parents, as compared with just under a quarter of middle-class students), more working-class students had come to this C.A.T. than would have gone to university. As one expressly put it, "they fitted better at a C.A.T." "University is beyond me and the background I was brought up in," said one student, whilst another regarded university as "on a lofty plane" and yet another held that "university is the tops but for other people." It was clear that, to some students and to their parents, there was a 'stigma' attached to a university—the word was used by one student—and university was regarded as 'snobbish and intellectual.'

Martin Trow, commenting on the Robbins' Report, wrote: "If the American experience has any relevance, the first experience of most working-class youth in institutions of higher education will not be in honours courses in universities, but in the schools and colleges where higher education is more directly linked to the world of work . . . The explanation is not very complicated and is well documented—it reflects differences in language, cultural resources, class attitudes towards education. These are simply facts of life." The survey supported these views. It was the link with the world of work, the practical and vocational orientation of the courses at the C.A.T. which made them acceptable to some students and their parents; indeed, one Diploma in Technology student indicated that his parents thought he was out at work, and another said: "they think this is part of the job." This aspect of the sandwich course, as a form of higher education which is socially acceptable to a boy from a working-class background who would not consider a full-time university course, has been inadequately appreciated in this country. To quote Trow again—"In so far as Britain aims to reduce class differences and democratize possession of and participation in the national culture, then it has to provide the kind of institutions that first generation students will find attractive and rewarding. Some working-class youths can make the jump to the university in one generation. But I suspect that for the bulk of able and talented people from working and lower middle classes, it will not be the university." The sandwich course may be providing just the bridge needed to enable many working class youngsters to enter a university institution in one generation.

V.—REFERENCES.

- CENTRAL ADVISORY COUNCIL FOR EDUCATION (ENGLAND) (1954). *Early Leaving*. London: H.M.S.O.
 FLOUDD, J., HALSEY, A. H., and MARTIN, F. M. (1957). *Social Class and Educational Opportunity*. London: Heinemann.

- CENTRAL ADVISORY COUNCIL FOR EDUCATION (ENGLAND): 15 to 18. London: H.M.S.O., Vol. 1, 1959, Vol. 2, 1960.
- JACKSON, B., and MARSDEN, D. (1962). *Education and the Working Class*. London: Routledge and Kegan Paul.
- KAHL, J. A. (Summer 1953). Educational and occupational aspirations of 'common man' boys. *Harvard Educ. Rev.*
- ALLEN, *et al.* (1962). *Student Social Survey*. Nottingham University.
- CENSUS (1951). *England and Wales, Occupation Tables*. London: H.M.S.O. 1956.
- CENSUS (1951). *Classification of Occupations*. London: H.M.S.O., 1956.
- ELLIS and CLAYTON-LANE (October, 1963). Structural supports for upward mobility. *Amer. Soc. Rev.*
- TROW, M. (March, 1964). Second thoughts on Robbins. *Univ. Quart.*, Vol, 18, No. 2.

(Manuscript received 7th November, 1964.)

CO-EDUCATION : THE VERDICT OF EXPERIENCE

II.—A QUALITATIVE APPROACH

By R. R. DALE

(*Department of Education, University College of Swansea*)

SUMMARY. In a previous article the writer reported an enquiry into the opinion of grammar school teachers in Glamorgan about co-education. This was centred on the experience of teachers who had taught in both co-educational and single-sex schools and was quantitative in type. The questionnaire used was open-ended, and an account of the qualitative results derived from the comments was promised in this later article. A great majority of the comments were in favour of co-education. These are classified, and several interesting differences between groups of teachers are indicated. Opinions against co-education are similarly dealt with.

I.—INTRODUCTION.

IN Volume XXV, Part III of this *Journal*, the writer gave an account of an investigation made in Glamorgan on the opinion of grammar school teachers about the merits and disadvantages of co-education at the secondary stage. Under the title 'Co-education : the Verdict of Experience,' the article presented a quantitative analysis of the opinions of 83 per cent. of the teachers, classified according to present school, teaching experience and education, and the type of co-education (or its opposite) which they favoured. Co-education was defined as Type A (normal), Type B (mixed school with mainly single-sex classes), and Type C (Dual schools), Type C was, however, regarded as being much more akin to single-sex schools than to co-education and was classified in this way. The quantitative results, including the verdict of those teachers who had taught in both single-sex and co-educational schools, were heavily in favour of co-education. At the time this was reported, an undertaking was given that the respondents' reasons for their preference would be presented in a later article. This article, here published, has been long delayed by the pressure of other research.

II.—PROCEDURE.

At the time the work was commenced, only Moreton's work (1946) in the same field, had been reported. His work (and, after the event, that of W. E. Davies (1950)), showed the over-riding importance of ensuring a high percentage of returns. To this end, the enquiry was made by means of a simple questionnaire focused around the question "Are you in favour of co-education in secondary schools?" followed by "If you are in favour, which of the above types do you prefer?" The types were carefully defined (see above), as was also the term 'secondary school.' Comments were invited at the end of the form, so that respondents would be able not only to give their reasons for their preference, but also to protest against any ambiguity, etc., that they might find. Another measure to secure a high percentage of returns was the administrative procedure used for distributing the questionnaire. First, the area was deliberately limited to Glamorgan, and the co-operation of the local education authority was secured. Then heads of schools were asked for their co-operation in distributing the questionnaires to their staffs and collecting them again in sealed envelopes. In this way a very high percentage (83) of replies was secured from the staffs of those schools who co-operated; it is also certain that some of those

who failed to reply did so merely through administrative failure, e.g., one head apologised for a 50 per cent. return, explaining he was retiring in a few days. Another head refused to issue forms to two young newcomers on the staff on the ground of their inexperience: they were in favour of co-education. They are not, of course, included in the tables. It was unlikely, therefore, that these non-replies were mainly due to the opposition of teachers to co-education, expressed as non-participation. In addition, the sealing of the replies, and their anonymity, enabled the teachers to give an honest opinion, without being unduly influenced by the views of the head or by understandable reluctance to appear disloyal to the school in which they were teaching. Replies were received from 550 teachers in 32 schools. Of the four schools which failed to reply, one was co-educational (the head was ill), two were for boys and one for girls. Their omission would produce only slight bias in the sample. The conservative overall return of 83 per cent. breaks down into 85 per cent. from co-educational, 86 per cent. from boys' and 77 per cent. from girls' schools.

III.—RESULTS.

TABLE 1
QUESTIONNAIRE COMMENTS—PRELIMINARY ANALYSIS.

	Forms returned	With Comments	Favourable to Co- education	Against Co- education	Un- decided
Boys' Schools.....	169	100	71	27	2
Girls' Schools.....	166	95	50	36	9
Co-educational Schools..	215	121	110	7	4
Total	550	316	231	70	15

From the above raw data we see that 59 per cent. of the teachers in boys' schools who replied also made a comment: for girls' schools the figure is 57 per cent. and for co-educational schools 56 per cent., making an overall percentage of 57. Of the comments from boys' schools 71 per cent. were favourable and 27 per cent. unfavourable; for girls' schools only 53 per cent. were favourable, 37 per cent. being against; for co-educational schools no less than 91 per cent. were favourable with 6 per cent. unfavourable.

The favourable comments are classified in Tables 2A and 2B. It must be stressed that 43 per cent. of the teachers failed to make a comment. Though we know from their questionnaires that a decided majority of these were in favour of co-education, we cannot be sure which favourable reason they would have put forward. This account is, therefore, an analysis of the comments of the substantial number of teachers who took the trouble to give reasons for their preference.

A.—COMMENTS IN FAVOUR OF CO-EDUCATION.

The comments, like the quantitative preferences expressed, were heavily in favour of co-education.* In classifying them they were split up into their

* Of teachers in co-educational schools 91 per cent. were in favour of co-education Types A or B, 84 per cent. preferring the former; 5 per cent. wanted single-sex or dual schools. Of those in boys' schools 61 per cent. were in favour of Types A and B, while 33 per cent. preferred dual or single-sex schools. Of women in girls' schools 54 per cent. were in favour of Types A and B, while 34 per cent. preferred dual or single-sex schools.

components parts, so that one respondent's reply might be entered under one, two or more headings, according to whether he mentioned say 'Social life,' or 'Educative effect of the sexes' or 'General Discipline.' Inevitably, some of these categories overlap and also interact. The reasons given are primarily social and educational; easily the first three are 'Educative effect of the two sexes on each other,' 'Preparation for social life,' and 'Natural boy-girl relationship.' Important second rank reasons are 'School atmosphere,' 'School Societies,' and 'Competition,' with 'General Discipline' (for the Men) not far behind. Receiving still less mention are 'Academic Standard,' 'Staff Outlook,' 'Staff-relations,' 'Complementary Attitudes and Interests' (of the two sexes) and 'Different Interests.'

TABLE 2A

COMMENTS BY GRAMMAR SCHOOL TEACHERS WHO WERE IN FAVOUR OF CO-EDUCATION.

Type of Comment	MEN				Totals
	Teaching in Boys' Schools with experience in		Teaching in Co-ed. Schools with experience in		
	Both co-ed. and Single-sex Schools	Boys' Schools only	Both co-ed. and Single-sex Schools	Co-ed. Schools only	
SOCIAL :					
Preparation for social life	26	21	45	25	117
Educative effect of the sexes on each other ..	36	30	39	28	133
Natural boy-girl relation-ship	12	9	47	25	93
School societies	7	9	19	11	46
EDUCATIONAL :					
Different attitudes and interests	—	2	7	—	9
Complementary attitudes and abilities	—	—	5	2	7
Competition	11	7	4	4	26
Academic standard	—	—	6	2	8
TONE :					
General discipline	10	2	17	6	35
School atmosphere	14	7	14	7	42
STAFF :					
Staff relations	—	—	6	—	6
Staff outlook	—	—	5	2	7
Staff-pupil relations	—	—	4	—	4
OTHERS	—	5	4	5	14
Totals	116	92	222	117	547

N.B.—Single mentions are omitted.

TABLE 2B

COMMENTS BY GRAMMAR SCHOOL TEACHERS WHO WERE IN FAVOUR OF CO-EDUCATION.

Type of Comment	WOMEN				Totals
	Teaching in Girls' Schools, with experience in		Teaching in Co-ed. Schools with experience in		
	Both co-ed. and Single-sex Schools	Girls' Schools only	Both co-ed. and Single-sex Schools	Co-ed. Schools only	
SOCIAL :					
Preparation for social life	10	15	24	14	63
Educative effect of the sexes on each other ..	8	10	20	21	59
Natural boy-girl relation-ship	21	11	23	21	76
School societies	6	7	4	4	21
EDUCATIONAL :					
Different attitudes, interests and abilities	—	—	3	2	5
Complementary attitudes and abilities	—	—	2	5	7
Competition	8	3	6	6	23
Academic standard	3	2	2	—	7
TONE :					
General discipline	—	—	3	—	3
School atmosphere	9	4	9	5	27
STAFF :					
Staff relations	3	—	2	3	8
Staff outlook	—	—	8	—	8
OTHERS	3	3	—	—	6
Totals	71	55	106	81	313

N.B.—Single mentions are omitted.

The reliability of the classification was tested by giving a colleague fifty replies to classify independently. Agreement on the gross categories of 'Favourable,' 'Neutral,' and 'Against,' was complete. Disagreement on the classification into 'Social,' 'Educational' and 'Tone,' etc., was about 2 per cent. Within the Social section itself, however, where the comments classified as 'Preparation for Social Life,' 'Educative effect of the sexes on each other,' and 'Natural Boy-Girl relationship' tend to cross these somewhat arbitrary boundaries, some 14 per cent. of comments were classified in neighbouring sub-sections.

Though there are some differences between men and women in the reasons for their favourable attitude, as seen in the distribution of their reasons between the five large sections, these were not large enough to be statistically significant.* It will be seen later, however, that there is an interesting difference, within one of the sections.

* Chi-square is 7.74 for 4 degrees of freedom.

The individual comments are illuminating, especially when one considers the experience behind them; some are presented below under the headings already given. They are selected mostly from those teachers who based their remarks on their experience of teaching in both types of school; exceptions are made only for a few comments which are humorous or particularly well phrased, or from heads. Throughout the article, wherever a comment is quoted from a teacher whose teaching experience has been limited to single-sex or co-educational schools, this is indicated.

(i) *Educative effect of the sexes on each other.*

The central theme of these replies has several 'motifs.' One is that each sex tends to exert an ameliorating effect on the behaviour of the other, diminishing the occurrence of those extreme forms of behaviour said to be characteristic of each sex—the boisterous rowdiness and rude uncouth conduct of some boys and the giggling, simpering and 'cattiness' of some girls. Each sex unconsciously behaves better in front of the other. Another 'motif' is that boys and girls get to know each other in the everyday life of school and classroom, and each learns how the opposite sex thinks, how it reacts, what its interests are, while both acquire a respect for the abilities of the other. Girls learn how to respond to the good-humoured teasing of the boys, and boys become acquainted with the mercurial unpredictability of the female! Though the '*naturalness of the boy-girl relationship*' was of such frequent mention that it is classified separately, it is so closely related to '*the educative effect of the sexes on each other*' that this type of comment is included under the present heading.

But let the comments speak for themselves. A headmaster of a boys' school who was educated in a boys' grammar school but taught in a mixed school for many years, wrote: "In my experience the system has worked well—boys and girls worked well together. A class composed half of girls makes discipline much easier. The girls tend to make the boys less 'barbarous,' the boys tend to make the girls less 'gushing.' Senior girls may form romantic ideas about junior male members of the staff occasionally, but rarely with any serious consequences: the sex relationship of boys and girls have, in my experience, been quite healthy." The headmaster of a co-educational school, who was educated in a boys' school, wrote: "The sexes have a good influence on one another—the boys become more refined, and since I have been head of this school (more than 12 years) I have not had a single case of immoral conduct." Another head thought "the pupils respond more readily to discipline and the little courtesies and consideration for others receive greater attention."

A male teacher who was educated in a single-sex school and was teaching in a boys' school, condensed his impression rather pithily: "In the mixed school social life was more balanced, manners infinitely better. There was a brightness that took away much of the drudgery from teaching. In single sex boys' schools one encounters patches of arrogant coarseness and overt prurience which are depressing!" Another man, who had taught only in a boys' school but was co-educated, summed up as follows: "My seven years as a pupil in a co-educational school were full and happy ones. Boys and girls mixed freely together in and out of school, the result being that the boys and girls quickly ceased to regard each other as unknown quantities. This fact alone should outweigh any others which may be put forward against that type of school." A games master in his forties, himself a product of a boys' school, was quite definite: "Boys from co-educational schools are quite as hard and robust in their rugby as those from boys' schools, but the former are invariably better mannered, more polite and courteous. Generally speaking, the co-educated

boys are better disciplined in the field, more prepared to accept referees' decisions without question, and less boisterous and rowdy." Another man of the same background wrote: "I have found a far healthier atmosphere to pervade the school where the basis is co-educational." He himself had been a pupil in a boys' school. There is some nostalgic longing in the reply from a co-educated man who had taught only in a boys' school. "The community life of my youth in a mixed school gave me a joy which I find lacking in a boys' grammar school." Of the same character is the remark of an elderly co-educated man: "Class contact shows them that both sexes have their weaknesses, failings and conceits, and so obviates the 'Halo' danger. There is a natural atmosphere in the classroom, together with colour."

A woman in her late fifties who was co-educated and was teaching in a single-sex school after some years in a mixed one, summed her impressions up as "Children grow up in a natural way, as in a family." A rather frank statement is made by a very young woman who had experience only of single-sex schools, "It enables each sex to know and learn from the point of view of the other, contributing to better mutual understanding between the sexes and avoiding the rather immature attitude of one sex to the other which is often found in the products of single-sex schools." (!) An elderly woman, co-educated, wrote: "We have seen the boys deteriorate in behaviour in this district since they have been in a separate building." A woman in her thirties, with single-sex education, said: "Boys and girls learn to accept one another as equal partners, each on his or her own merit and yet to appreciate the special gifts and interests of the other sex. This is vital in the modern world." Another woman of the same background commented: "The contact is stimulating and healthy, and the sexes are much more unselfconscious in relationship one to another than when brought up in segregation." The following remark, from a young woman, gains in interest because of her single-sex education: "Co-education means that boys and girls are given the opportunity to be normal in their relationship with each other. They are so accustomed to working side by side as well as meeting each other at social functions that when they leave school and begin to work or attend college, they do not have to cope with additional difficulties."

This comment, like many others, begins with natural boy-girl relationships and finishes with preparation for adult life. A woman in her forties, educated in both types, believed that "Co-education promotes a more healthy attitude altogether among the sexes. Girls become less petty; boys become better mannered." Lastly is an amusing observation from a woman in her forties who had been only in girls' schools: "I feel that my experience is very limited, but theoretically at any rate if boys and girls are educated together they will behave more naturally in each other's company later on, and spend less time 'hunting' during their school life. It should also be of great advantage to the only child and very small or one-sex families."

Both sexes also made fairly frequent references to the companionship of the sexes in school societies, school athletics and games and the happy friendliness of school socials and dances.

(ii) *Preparation for life.*

The central theme is that the world has two sexes, and an education which keeps these sexes apart is not preparing children for life in the world.

A headmaster of a co-educational school condensed this theme into: "Co-education is the setting for maturing into healthy adult living." An ex-head who couldn't resist making an unauthorised contribution, was especially

interesting : " After 18 years in a boys' school, I became head of a large mixed school. My conversion to co-education was immediate and easy. I have no doubt that a better relationship exists among the pupils." A head of a boys' school, himself educated in a boys' school, was " sure that adolescents grow up healthier and better citizens from co-educational schools. Readjustment at College is also very difficult for pupils from single-sex schools."

A male teacher in his late thirties, himself educated in a boys' school hadn't the slightest doubt : " I am a whole-hearted advocate of co-education on educational, medical, psychological and social grounds. A co-educational school is the best preparation for a co-educational world." Another, co-educated wrote : " Education is for living in a *real* world, not in some hypothetical world of segregated sexes." Again, " The school should train children for living in the community and should be a microcosm of the whole community." An elderly gentleman, co-educated, had this to say : " The art of living in society is the art of living understandingly together. Life at home is the living together of father and mother, sons and daughters. A co-educational school is an extension of life at home and can serve as an apprenticeship to becoming an understanding member of society."

A younger man, educated in a boys' school, who may have been thinking of himself, wrote : " When mixing together in school boys and girls get to know a lot about each other. I consider this a great help when the question of choosing a mate has to be considered in later life." One of the best literary efforts was that of a co-educated man who had taught only in boys' schools : " The best type of school is that which is a microcosm incorporating within it the essential features of life in the world outside. Segregation of the sexes is purely artificial. Education cannot claim to be a training for life unless it prepares the child to take his place naturally in the community of men and women."

The remarks made by the women are similar. One, educated in a girls' school, stated : " I do not believe that any training in schools where the sexes are segregated can adapt children for an adult life in a world where there are two sexes. A girls' school affects girls in two ways, either they develop a rabid interest in sex or they withdraw into an idealised world." The comment of a middle-aged woman from a co-educational school was uninhibited : " At college, girls from the boarding schools behaved as if they were ' let loose '." More moderate was the following, from a teacher in a girls' school who had been co-educated : " I believe that boys and girls should be educated together, because when school days are over they will inevitably come in contact with each other, not only socially, but also in their daily work." The observation " Co-education gives a better training for life. I consider segregation of the sexes unhealthy during the formative years," takes on a new interest when one notices that the teacher making it was herself educated in a single-sex school. An elderly woman who had not taught in a co-educational school, but was brought up in one, declared : " Co-education is a good preparation for both sexes to live together and co-operate in after school life. There is less tendency for either sex to ' break out ' (than) after the regime of a single-sex education. The outlook is far healthier."

(iii) *General discipline and school atmosphere.*

These two categories are sufficiently close together to be amalgamated here. Comments on general discipline maintain that the discipline of a co-educational school tends to be better than that of a boys' school because the

presence of the girls makes the boys less boisterous. Those on school atmosphere stress the naturalness, friendliness and 'family' impression of the tone of the school. The men place marked but almost equal stress on discipline and atmosphere; these comments are based on comparisons between mixed and boys' schools. The women, on the other hand, stress only atmosphere (8.6 per cent. of favourable comments), and make little mention of discipline (just below 1 per cent.)*; they are comparing mixed with girls' schools. Typical comments were:

The head of a co-educational school, who was educated at a boys' school, stated: "The corporate activities of the school are greatly enriched by the presence of the two sexes." Another who had had long experience in boys' schools wrote: "A better relationship exists between the staff, and between pupils and staff, in a mixed school." A third head of a co-educational school who was educated at a boys' school observed: "A far better spirit is engendered in the school and attitude to work is better."

The remark, "The standard of behaviour of boys in co-educational schools is higher than in boys' schools as such," written by an elderly co-educated man, is typical of many replies, as is the more colourful "Co-education softens a boy's 'barbaric' traits, and gives girls greater confidence." Typical of the second aspect in this section is, "Co-education usually results in a friendlier and more homely atmosphere in the school. In addition, it has a beneficial effect upon the manners and behaviour of both sexes, thus imparting a better 'tone' to the school." A third man, co-educated, wrote: "The presence of girls has a marked effect on the appearance, conduct and work of the boys. A co-ed. school has a brighter, less monotonous atmosphere."

People do not like to appear to be disloyal to the school in which they have been educated, but one woman educated in a girls' school admitted: "There seems to be a healthier atmosphere among the children (in a co-educational school) and a greater sense of sportsmanship." The same point applies to the following: "Pupils are easier to manage—boys less noisy and rough, girls less self-conscious and less inclined to giggle or form cliques; they appear to have a moderating influence on each other. Both girls and boys develop a more natural ease of manner and quiet assurance."

(iv) *Educational reasons.*

The main arguments put forward under this heading are the friendly spirit of competition between the sexes in the classroom, and the beneficial influence and broadening effect of the greater breadth of interest. The former reason was expressed by a middle-aged man whose only experience had been in co-educational schools: "In school work co-education caters for an inspiring spirit of rivalry and competition." The second is exemplified in the remark of a man who was educated in a boys' school: "Since boys are generally more practical and girls more literary, they exert a beneficial influence on each other. I noticed an all-round improvement in these respects when a boys' school in which I taught was combined with a neighbouring girls' school." This opinion was echoed by an elderly man who had no experience of boys' schools: "Boys' interests and background of knowledge differ from that of girls. Teachers cater for these and make use of them in their lessons." Another man who was educated in a co-educational school, but had never taught in one, stressed the broader interest: "The presence of girls in the Arts classes provided a natural atmosphere for the studies. Dramatics were most enjoyable as mixed efforts."

* The difference between the comments of men and women is statistically significant beyond the .001 level (Chi-square test),

On the other hand, an elderly male thought that "Boys tend to raise the standard of the girls in certain subjects, e.g., Mathematics and Science." Another man played the same tune in a different key: "Competition between the sexes produces better work. The girls up to 15 plus are more conscientious workers and lead the boys to a higher standard of achievement." Finally, a man of wide experience stated: "As a Mathematics teacher, I find boys are naturally quicker at working out mathematical problems and they help to create an enthusiasm for the subject. On the other hand, girls are more conscientious in working out the longer but easier questions and set a good example of neatness and perseverance."

The comments of the women were not very different, though rather fewer of them were favourable. One middle-aged woman, educated in a girls' school, wrote: "I consider that co-education is a more natural system—the contribution of opinions and ideas on all subjects are more varied and wider in range than is possible in a single-sex school." A fairly general comment comes from another woman of similar experience: "The spirit in the (co-educational) classroom is a much more lively one and the teaching of Dramatics is far easier and more interesting." Two women, both educated in co-educational and girls' schools, but only the second of whom had taught in a co-educational school, wrote as follows: "For about 15 years, I have taught VIth Form boys (from another school), with VIth Form girls, in Biology, and it has worked well. Boys are bolder and show more initiative in work, girls are more thorough; each influences the other. There is a healthy rivalry in work. Boys' manners improve and there is a pleasant friendly feeling throughout classes without any silliness." And "In my own subject—Chemistry—boys are, on the whole better than girls, I think, and have a good influence on the class—while I think the girls have something to teach the boys in their care of detail in practical." Mentioned several times was the idea that teaching in a co-educational school was "more interesting from the teacher's point of view, as boys often express opinions more spontaneously than girls; this is especially helpful in VIth Form work."

(v) *Staff.*

Remarks in this category were more frequently from the women than the men. One middle-aged teacher observed, "In a co-ed. school the emphasis tends to be on essentials rather than on details and red tape." Another wrote: "I was happier and the other women appeared happier in co-ed. schools. Therefore, I should always wish, if circumstances permitted, to teach in one." Repeating the same theme was, "Single-sex schools can have a very narrow-minded viewpoint among the staff." This was the opinion of a woman of 40 who was educated in a girls' school and had taught in both girls' and co-educational schools.

(vi) *Other.*

One schoolmaster in a senior position who had long experience of boys' and co-educational schools, maintained that "Opposition to co-education comes invariably from Headmistresses for economic and psychological . . . reasons."

B.—COMMENTS AGAINST CO-EDUCATION.

We now come to the comments against co-education, which were much less numerous.

TABLE 3A

COMMENTS BY GRAMMAR SCHOOL TEACHERS WHO WERE AGAINST CO-OEDUCATION (MEN).

Type of Comment	MEN				Totals
	Teaching in Boys' Schools with experience in		Teaching in Co-ed. Schools with experience in		
	Both co-ed. and Single-sex Schools	Boys' Schools only	Both coed. and Single-sex Schools	Co-ed. Schools only	
SOCIAL :					
Sex dangers	1	3	—	—	4
Boy-girl relationship	3	4	3	1	11
EDUCATIONAL :					
Different rates of intellectual development ..	—	—	1	—	1
Different attitudes and interests	1	6	5	3	15
Competition	—	2	1	—	3
Academic standard	4	4	5	2	15
TONE :					
School atmosphere	—	—	1	—	1
General Discipline	2	8	2	1	13
STAFF :					
Pupils and teacher should be of same sex	—	7	3	—	10
Staff-pupil relations	1	2	1	—	4
ORGANISATION :					
Organisation, including unfavourable treatment of girls	1	3	5	3	12
Promotion prospects	—	—	—	—	—
OTHERS	1	2	5	1	9
Totals	14	41	32	11	98

Comparing this table with the table of favourable comments we note that, whereas the latter was heavily weighted in the social categories without the other categories being neglected, the unfavourable comments are concentrated on the educational and organisational sides.* We must keep in mind, however, the possibility that some teachers were reluctant to put on paper—even though anonymously—their fears that co-education might cause sexual misconduct. Here, however, we can only detail the thoughts expressed. The number of comments in each section is admittedly small, and these are classified into sub-sections only to show that they contain no surprises, and to keep the format parallel with that of Table 2.

* E.g., the difference between the distribution of the favourable and unfavourable comments, for women, is statistically significant well beyond the .001 level.

TABLE 3B

COMMENTS BY GRAMMAR SCHOOL TEACHERS WHO WERE AGAINST CO-EDUCATION (WOMEN).

Type of Comment	WOMEN				Totals
	Teaching in Girls' Schools with experience in		Teaching in Co-ed. Schools with experience in		
	Both co-ed. and Single-sex Schools	Girls' Schools only	Both co-ed. and Single-sex Schools	Co-ed. Schools only	
SOCIAL :					
Sex dangers	—	—	—	—	11
Boy-girl relationship....	7	4			
EDUCATIONAL :					
Different rates of intellectual development ..	4	4	2	2	12
Different attitudes and interests	3	6	1	—	10
Competition	—	1	—	—	1
Academic	3	3	2	1	9
TONE :					
School atmosphere	1	1	—	—	2
General discipline	4	2	1	3	10
STAFF :					
Pupils and teacher should be of same sex	1	2	—	1	4
Staff-pupil relations	1	1	—	1	3
ORGANISATION :					
Organisation, including unfavourable treatment of girls	11	4	3	1	19
Promotion prospects for women	—	10	1	—	11
OTHERS	1	1	1	1	4
Totals	36	39	11	10	96

(1) *Social (boy-girl relationship).*

Only four men and not one woman, teaching in co-educational schools, voiced any doubts under this heading, though four men and seven women who had taught previously in such schools did so. One man, aged 50, with no experience of co-educational schools, believed that, "Between the ages of 5 (approximately) and 15 (approximately) there is no desire on the part of either sex to associate with the other sex." A woman of the same age and background wrote: "For girls co-education is a disadvantage as they would be superseded in class by the boys. They would also lose much refinement." Another woman whose only contact with co-education was as a pupil, stated: "Co-education leads to general slackness and untidiness and a general disregard for the finer points of behaviour, good manners and even dress."

The four men teaching in co-educational schools, mentioned above, wrote as follows : " Girls 16+ have an unsettling effect on boys." " Tone of school suffers compared with types where boys and girls are separate. Believe it due to a difference in attitude to discipline, especially younger boys who seem to dislike girls and their ' code '. Would suggest 15 onwards (full co-education) is definitely the best." " Instead of girls having a refining influence on the boys, the tendency is for boys to become a little familiar and coarse in their attitude. I feel they do not show enough respect for the ' fair sex '." " If co-education makes for easier relations between the sexes, it also (as far as I have observed) produces less independence of mind, less virility. Discipline though easier with co-education produces less real self-discipline. At least that appears to be true of the conditions I have experienced—the reasons for the disadvantages may possibly lie elsewhere." The last man had taught in only a co-educational school ; he was, however, educated in a boys' school.

(ii) *Educational.*

Sub-sections under this heading comprise the different rates of intellectual development of the sexes, the different interests, different abilities, boys and girls distracting each other from work, and the bad effect of competition on the girls. A headmistress combined several of these in " The girl develops emotionally (biologically) earlier than the boy—the outlook and interests of the two sexes from 11-16 years are very different. In many cases, the educational progress of girls (13-15) is affected in a mixed class—either the girl becomes very ' silly ' and pre-occupied with boys or very self-conscious in the presence of boys and will not take an active part in the lessons." Several comments resembled the following, written by a man who had no co-educational experience, " The scholastic standards of single-sex schools are higher generally than those of the mixed school." There is more flavour in " The social graces allegedly acquired in a co-educational school by boys seem to be acquired at the expense of their work. The thesis that girls would inspire boys with a spirit of emulation is doubtful—the only competition ensuing would be for the girls' favours." This lady had no experience of co-education. A man with the same restricted background believed that, " Co-education provides distraction during the day ! " Another man supported this : " In certain cases co-education has harmful effects—the opposite sex exerting a distracting influence on the other." A co-educated woman with no co-educational teaching experience maintained that " A relationship (of teacher to class) that would be natural and unconscious with one sex becomes artificial and self-conscious when the sexes are mixed. Younger teachers who do not always realise how differently the sexes react, often irritate and alienate one-half of the class by a manner and method which would be entirely successful with the other half alone." Similarly, a man with no co-educational experience affirmed : " Boys and girls of the same age are at different stages of social development and some subjects can't be presented in the same way for both—e.g., in classics boys of 13 enjoy the direct method while girls of the same age are too ' grown-up ' and self-conscious." A male supporter of co-education had some doubts on the educational side : " Though the academic achievement may be lower . . . the social advantages more than counterbalance."

(iii) *School atmosphere and discipline.*

The principal argument here is that the presence of boys has a bad effect on the girls, though one man who had no experience of co-education thought : " Boys are hard enough to handle without adding the further distraction of

girls," and another believed that "the atmosphere created by an all-men staff in an 'all-boys' school is probably the best atmosphere in which to train a boy for real service in after school life." This man, though co-educated, had never taught in a mixed school. A woman with no experience of co-education commented: "In schools of one sex discipline is better and pupils are better able to concentrate on their work."

(iv) *Staff.*

In this section appear few but very trenchant comments. Some illustrate the prejudice and high emotional feelings which may still be found. Short and sharp is, "Men should teach boys and women girls," a comment made by a man in his sixties who had no co-educational experience. More hard hitting is: "The chief reason against co-education is the poor quality of staff obtainable. The only men who are likely to apply for jobs in mixed schools are those who are unable to hold down a job in a boys' school; the women are those whose chief aim is to get a husband of some kind. Neither of these is likely to be efficient or to put the children's interests first." She did not record whether she herself had taught in a co-educational school! A male teacher, who had not put the matter to the test, maintained that "Life is easier and less distracting, both in class and staff-room, if one sex only is involved." One of the most scathing comments was by a woman who was quite unaware that she herself was included in the condemnation: "I have had considerable experience as a teacher and pupil in co-ed. schools. My objection to the system is based on the indubitable fact that only the poorest type of staff can be recruited for these schools."

(v) *Organization.*

The feminine opponents of co-education—and some of its feminine supporters—find two objections under this heading, the lack of promotion for women in co-educational schools, and a tendency for girls to be neglected in favour of the boys. A point never put forward as a major factor is the greater difficulty in organising a co-educational school. No men made comments in this section. Typical of the first objection is: "One aspect of co-education that seems to me unfair is that in such a school there is less prospect of promotion for women teachers." This comment occurs several times. A headmistress, experienced in both types of school, believes that "Health problems and the general welfare of the girls does not generally receive the same attention in a mixed school . . . as in a girls' school." A possibly justifiable complaint was: "My experience of co-education suggest that the boys are well catered for in the curriculum at the expense of the girls. In the school which I attended, Botany was the only science subject available to girls." The school was co-educational and she had taught in both types of school. Another woman with similar experience wrote: "The type of staffing and the selection of subjects are studied from the boys' angle, and good girls are often denied study in a particular subject (e.g., Mathematics)."

IV.—DISCUSSION.

The study of a part of a community in isolation is especially dangerous when the aspect studied concerns beliefs and attitudes. The opinions of the members of the part will inevitably be affected by the *mores* and prejudices of the larger community; these people are, indeed, part of this community and cannot separate themselves from it. Schools are no exception to the rule; they are a reflection of the society in which they are placed, and of the society

in which they were placed in the past. The reflection is tinged with idealism because the school attempts to teach its pupils to behave not as society itself behaves, but in accordance with society's ideals.

The decision to organise schools co-educationally or as single-sex institutions comes from society's *mores* and beliefs, and also its economic limitations. In Victorian times, when it was proper to be prim, and young ladies were chaperoned even from dawn to dusk, and later, society's fears in relation to sex would have powerfully affected the opinions of teachers about co-education. Today, though these fears are still expressed in a few quarters, and are maybe only dormant in some others, they no longer have the overwhelming influence they formerly had. None-the-less, most of the opinions on co-education reported in this article, though reliable for the present age, might conceivably change with any profound change in the *mores* and beliefs of society. The qualifying phrase 'most of' is inserted because the question of the comparative standards of attainment in single-sex and co-educational grammar schools has recently been largely removed from the realm of opinion to the realm of fact by the work of a number of researchers (Dale, 1962 and 1964).

Nor is this theory of co-education necessarily applicable with equal force to all communities, or to different periods within the history of one community. Its success depends in part upon the support it receives, or the handicaps under which it labours, deriving not only from the ideals but also the practices of the society around it. Again, even if we were to suppose that the recorded opinions of a large majority of those teachers in this research, in favour of co-education in grammar schools, were decisive (though this is not claimed), and that the verdict is valid for the generality of children, no one could say that it was valid for all children or even for all grammar school children. Temperaments vary, and though most children might be well suited by co-education, it is not impossible that some individuals might be better suited to single-sex schools. The same argument applies, perhaps with rather greater force, to teachers. We must also remember that within each group of schools, whether single-sex or co-educational, there will be a wide range of 'educational efficiency.'

To what extent would the picture be similar in other parts of England and Wales? Returns are available for only eight additional schools in Wales, but the findings are the same. The few schools in England which made returns pointed in the same direction with probably some diminution in the majority in favour of co-education, though the evidence is too thin for a definite finding. It can, however, be said that in all co-educational schools investigated there was a near unanimity in favour of co-education, whereas in single-sex schools there was almost invariably only a minority in favour of their own regime. The English public schools were, however, not investigated. The findings of F. Moreton (1946) and W. E. Davies (1950) strike the same note.

Now to turn briefly to the comments themselves. The support given to co-education by these teachers is based on educational reasons, interpreting education in the fuller sense to include social and emotional development. This is also the verdict of those teachers who had teaching experience in both single-sex and co-educational schools. Though they stressed the social aspects they also gave a prominent place to school atmosphere and discipline, and did not neglect the instructional side.

Not one of the teachers then in co-educational schools reported adversely on 'sex dangers' in such schools, nor did any woman in the whole survey; such objections were limited to four men, three of whom had no teaching experience in co-educational schools. Opinions against co-education were mostly based on instructional grounds, including distraction of one sex by the

other and an allegedly lower academic standard due, it is implied, to co-education. As we now know this allegation is not true in the case of boys and probably not true in the case of girls, these objections should gradually vanish.* The next most important class of objection, that based on organisation, could be met in two ways. The lack of promotion for women in co-educational schools might be remedied—at least in part—by giving the senior mistresses the title of headmistress and rather more responsibility; she should have seniority in rank above the senior master. The second objection—the favourable treatment of boys at the expense of girls—would be partly met by the increase in the status of the headmistress and partly by the male heads being made aware of this possible bias. The problem may occasionally arise when a head is appointed who has had no previous experience of co-educational schools. Opinion, even if based on experience, cannot be conclusive. The search continues for more objective measures which will provide factual evidence and thereby raise the discussion on to a new plane.

* See Clark (1937), Dale (1961, 1962, 1964), Gott (1938), King (1949), Sutherland (1961), Tyson (1928) and Walton (1934).

ACKNOWLEDGEMENTS.—The writer acknowledges with gratitude the financial aid given by the Department of Education and Science for his research in this field. Dr. Philip Williams is thanked for kindly reading the manuscript and making some useful suggestions. Thanks are also expressed to Glamorgan Education Authority and the heads and staffs of Glamorgan grammar schools for their co-operation, and to my secretary, Mrs. R. Lewis, for her indefatigable assistance.

V.—REFERENCES.

- CLARK, G. (1937). *Co-education—Analysis of the Work and Principles of the Mixed School*. Unpublished M.A. thesis, University of Liverpool.
- DALE, R. R. (1955). Co-education: the verdict of experience. *Brit. J. Educ. Psychol.*, 25, III.
- DALE, R. R. (1948). Article in *Times Educational Supplement*, August 28th.
- DALE, R. R. (1949). Article in *A.M.A.* (September).
- DALE, R. R. (1962). Co-education: a critical analysis of research on the effects of co-education on academic attainment in grammar schools. *Educ. Res.*, 4, 3.
- DALE, R. R. (1962). Co-education: an analysis of research on comparative attainment in Mathematics in single-sex and co-educational grammar schools. *Educ. Res.*, 5, 1.
- DALE, R. R. (1964). Co-education: research on comparative attainment in English in single-sex and co-educational grammar schools. *Educ. Res.*, 6, 3.
- DAVIES, W. E. (1950). *A Study of the Attitudes of Secondary School Teachers Towards Co-education, Differentiation of Curriculum and Sex Teaching*. Ph.D. thesis, University of London.
- GOTT, B. (1928). Investigation while Director of Education for Middlesex. Details not available, except table of statistics in HOWARD, B. J. (1928), *The Mixed School*, London.
- KING, W. H. (1949). *A Critical Analysis of the Results of a School Certificate Examination*. Unpublished Ph.D. thesis, University of London, Institute of Education.
- MORETON, F. E. (1946). Co-education—a statistical enquiry into the attitude of teachers towards co-education, and a comparative study of the emotional development of children trained in co-educational and other institutions. *Brit. J. Educ. Psychol.*, June, 1946. (Summary of Ph.D. thesis, University of London, 1939.)
- SUTHERLAND, M. (1961). Co-education and school attainment. *Brit. J. Educ. Psychol.*, 31, 2.
- TYSON, G. (1928). *Some Apparent Effects of Co-education Suggested by a Statistical Investigation of Examination Results*. Unpublished M.Ed. thesis, University of Manchester.
- WALTON, H. M. (c.1934). Unpublished typescript. Middlesex Education Committee.

(Manuscript received 11th November, 1964.)

THE STRUCTURE OF PERCEPTION: A FACTORIAL STUDY

By R. M. BEARD

(Education Department, University of Birmingham)

SUMMARY. A factorial study of twenty tests of visual and auditory perceptions and memories, and ratings in mathematics and English for boys and girls aged about 15, gave rise to seven factors: reasoning with verbal abilities, five perceptual factors and a factor contrasting an easy perceptual task with ratings in school subjects.

The perceptual factors were (1) speed and strength of closure, (2) a spatial factor concerned with direction and orientation, (3) recognition and recall of Gestalten, (4) flexibility of closure, and (5) perceptual span. Almost identical factors appeared for both sexes and no perceptual differences were found for different kinds of school attainment, possibly owing to the general nature of the perceptual factors obtained. It is suggested that these factors have a general nature due to the way in which the brain functions, through the generalising power of language and, possibly, through kinaesthetic cues which link auditory and visual perceptions.

I.—INTRODUCTION.

ALTHOUGH many studies have been made of perceptual abilities, the majority have been concerned with visual perceptions; there have been few attempts to see in what ways various kinds of perceptions are related or to determine the part played by perceptual abilities in different kinds of intellectual skills. Thurstone⁶ considered the possibility that factors found in tests of visual perceptions might also be found in other perceptual fields. But, although he and his students related these visual perceptions with cognitive abilities and with personality traits, as well as with abilities in reading and in several occupations, they did not follow up his speculation that the factors of perceptual closure found with visual materials might transcend the modalities.

The primary aim of this investigation was to see whether Thurstone's closure factors, and any other perceptual factors, obtained with visual materials, arose also from auditory or kinaesthetic tests. Secondary aims were to relate perceptual abilities and different kinds of memory with mathematical and literary abilities and to discover any sex differences in the nature of the factors.

Subjects were selected from the whole range of ability among school children of about 15 years, so far as this was possible. For the purpose of identifying factors this had the disadvantage that intelligence might play so large a part in success that group factors would be less pronounced, or indistinct, unless a very large sample was taken; on the other hand, it had the advantage that group factors obtained would not be peculiar to a group of persons with some special kinds of experience. The children came from five mixed schools: a grammar school, a grammar school stream in a comprehensive school, a commercial-technical school, and two modern schools, one from a fairly prosperous working class suburb and the other from a city slum. Since only the grammar school children were selective at the age of 11 and were assigned to this school because they then had I.Q.s. of 120 or less, it is probable that the highest range of intelligence is rather thinly represented, but many of these children were to take O-level or some other quite difficult examination so it is possible that those

of moderately good intelligence were too numerous in the sample. However, in Heim's A.H.4 test of verbal intelligence, the majority of the 145 scores lie symmetrically between 26 and 50; there are no scores above 50 but there is a tail of ten scores in the range 6-25. Children of very poor intelligence were well represented owing to the practice in the slum school of keeping children rated E.S.N. among the more able. In the corresponding non-verbal test the scores form an almost symmetrical distribution between 21 and 65 but four children obtained scores between 16 and 20.

II.—THE TESTS.

There are twenty tests and two sets of ratings in the battery described here. These remain after exclusion of relatively unreliable tests of kinaesthetic perceptions and of personality test scores.

Title of Test.

- | | |
|---|--|
| 1. Street Gestalt Completion. | 30 incomplete pictures duplicated on a single sheet. As many as possible to be recognized in 3 minutes. |
| 2. Dotted outlines. | 16 dotted outlines of letters and numbers. Each exhibited on a large card for 3 seconds. |
| 3. Mutilated words. | 12 mutilated words, somewhat easier for school children to recognize than those used by Thurstone. Each exhibited for 4 seconds. |
| 4. Completion of common exclamations. | 30 common exclamations to be completed. As many as possible in 2 minutes. |
| 5. Word pairs. | 30 incomplete pairs, as many as possible to be completed in 2 minutes. |
| 6. Descriptive phrases. | 30 incomplete phrases, as many as possible to be completed in 2½ minutes. |
| 7. Mazes A. } | Group versions of a new perceptual maze test, described below. Scores were number of mazes solved in 15 minutes for mazes A and in 10 minutes for mazes B. |
| 8. Mazes B. } | |
| 9. Gottschaldt test, page 1. | 30 simple geometrical figures to be found in 30 more complex ones. 2 minutes allowed. |
| 10. Gottschaldt test, page 2. | A corner shape to be found in seven successive figures. One of two symmetrical figures to be found in 7 more complex figures. 2 minutes allowed. |
| 11. Gottschaldt test, page 3 | Two pairs of figures, one or other of each pair to be found in sets of ten successive intricate figures. 3 minutes allowed. |
| 12. Heim A.H.4 test verbal reasoning. | 65 items comprising verbal opposites, synonyms, analogies, numerical sequences to be continued, and operations with numbers following written instructions. 10 minutes allowed after practice. |
| 13. Heim A.H.4 test non-verbal reasoning. | 65 sets of geometrical figures to find analogies, identities, results of combining or separating figures and completion of sequences of figures. |
| 14. Memory for digits. | Sets of six and nine digits to be remembered in order. Sets of five and seven digits to be remembered and recorded in reverse order. |
| 15. Memory for sentences. | Terman-Merrill L-test. Average Adult 7 (b). The first three sentences of M-test age XIII, item 2. Scored 0, 1 or 2 for each phrase. |
| 16. Memory for designs. | Terman-Merrill L-test, age IX, item 3 (a) and (b). In addition, an unsymmetrical design of three wavy lines with two scrolls. Scored 1, 0 or 2 for certain relationships in each figures. |

- | | |
|---------------------------------------|--|
| 17. Paper folding. | Terman-Merrill L-test, age IX, item 1 (b); Superior Adult II item 4; also a square folded diagonally, the corners folded to the right angle giving a smaller square, a square hole cut from the centre. Scored for correct number of creases and holes, and relationships of size. |
| 18. Memory for rhythm. | Every fourth item of Seashore's rest, recorded on tape. |
| 19. Wing's test of musical ability I. | Analysis of chords. The number of notes to be stated in each of twenty chords. |
| 20. Wing's Test II. | Pitch discrimination. Any change in a note of 30 repeated chords to be recorded: "up," "down" or "same." |
| 21. Rating in English | Ratings on fifteen point scales were obtained for each subject. |
| 22. Rating in mathematics. | |

In test 4 an attempt was made to provide a verbal exercise less likely to favour children from high status homes than is usually the case. Catch phrases and exclamations, likely to be heard on television or in daily use, were to be completed. However, it proved impracticable at the time to put these onto tape so the test is a speeded paper and pencil test like tests 5 and 6.

Tests 7 and 8 are relatively new tests, not previously included in a battery of this kind with children. They are group versions of a perceptual maze test devised by Elithorn, Kerr and Mott.² In each maze subjects are required to find a path through the maximum number of dots arranged at the intersections of lattices within a rectangular configuration. The authors claim that once the method has been learned no further logical reasoning is needed to obtain solutions but that success is determined mainly by a perceptual skill which may be a perceptual scanning mechanism. It seemed probable that with subjects so poor intellectually as some in this investigation, ability to count and to recognize and combine groups of dots accurately would also be important.

The three tests numbered 9 to 11 usually appear as one test with older subjects. During pilot testing with a group of boys they were found to respond very differently to three sections of the Gottschaldt test. This appeared to be due to the relative ease of the first page as compared with the third one and to the greater possibility of errors in orientation on page 2. On page 2 the choice of a corner shape in successive diagrams frequently showed indifference to its orientation and sometimes to the size of the unit; the succession of different shapes on page 1 and the more complex shapes of the latter part of page 2 and page 3 did not give rise to this kind of error. In young children errors in orientation are normal but it may be that this kind of error is more common among school leavers than is usually supposed. The tests were included primarily as tests of flexibility of closure, but, as we have seen, test 10 has a spatial component.

Ratings on a fifteen-point scale were obtained for ability in English (21) and in mathematics (22). Points on the scale ranged from A+ (brilliant, original, might get a first class degree) to E- (appears to have no number concepts or cannot read or write). The extreme ratings were not used by the teachers, their ratings in English ranged from A- (could take Advanced level) to E+ (very weak, several years retarded relative to his age group). In mathematics, the extreme ratings given were a point lower, from B+ (should get a good O-level pass) to E (can use only the simplest number combinations).

III.—ANALYSIS OF THE DATA.

Mean scores and standard deviations in the tests were calculated for each sex separately (Table 1). Mean scores for the girls tend to fall below those of

the boys in tests of visual perception and are considerably lower than those of the boys in the test of non-verbal reasoning. Mean scores in verbal reasoning, however, are almost equal for the sexes. It is possible that in this sample the girls are a little less able than the boys but the considerable discrepancy between the means for the two tests of reasoning suggests rather than there is a difference of experience or interest which accounts for the superiority of boys with visual materials.

Distribution of scores in the tests do not deviate to a considerable degree from the normal and, where they deviate, tend in the expected direction of being somewhat negatively skewed.

Correlation matrices were calculated for each sex separately. In Table 1 correlations for the girls appear as an upper right-hand triangular matrix, while those for the boys form a lower left-hand triangular matrix within the square matrix. Mean correlations are shown for each sex, and are appreciably higher for boys. A similar disparity was found by Withrington¹⁰ who included tests of memory and imagery with tests of mathematical ability. He found that correlations between tests were generally greater for boys, being much greater for spatial tests; girls, however, excelled in verbal intelligence, and memory tests were better done by verbalisers. In this study correlations are great on the whole for boys except between verbal tests, but the somewhat greater variability of the boys, as shown by their larger standard deviations in most of the tests, may contribute partly to the larger correlation.

The correlation matrices for the twenty-two tests were subjected to Principal Component factor analyses, seven factors being extracted in each case. These were rotated by Varimax to approximate orthogonal simple structure and further slight rotations were made to reduce the number of negative loadings and to give better psychological meaning to the factors. For the girls an orthonormal solution was satisfactory but for the boys an oblique solution might have given better results as some of the factors obtained were correlated. However, the correlations were not high except between factors A and C.

Before describing the rotated factors, justifications for extracting so many will be briefly outlined; a recent paper which puts the case for the use of minor components, arising from Principal Component analyses, makes a longer discussion unnecessary. Slater⁵ points out that there are no general grounds for assuming that dimensions where variation is restricted are uninteresting; and he shows that minor components may be stable, reliable and have psychological importance. For example, a test re-test study of a battery of twelve tests proved the eleventh component to be as reliable as the second. Further, Slater shows that possible tests for the stability of components are their stability under repeated sampling and the appearance of the same component vectors after the introduction of new tests.

In this study, since few subjects were used in each analysis, errors are fairly considerable; we have not attempted to extract twenty-two components but have used seven, extracting components until the variance accounted for by successive ones ceased to diminish appreciably. The fairly close similarity of the components in the analyses for boys and girls, and the stability of both when further analyses were made, including the ratings in mathematics and English, suggest that they comply fairly well with the two tests for stability. In addition, five of the components, including the seventh, can be identified as factors found in other batteries of perceptual and cognitive tests.

Although Slater rejects the word 'factor' in connection with principal component analyses, preferring the word 'component,' we have followed the

more usual practice of talking of 'factors' unless, as in this discussion, it would result in confusion.

In the description of the factors which follows, all loadings exceeding 0.2 are listed; but in interpreting the nature of the factors only loadings of 0.4 or more are used. The smaller loadings are of interest in confirmation of the nature of the factor if they appear for both sexes, or they may highlight differences between the sexes where certain tests have appreciable loadings for one sex only.

IV.—THE FACTORS.

Test	Factor A	
	Boys	Girls
12 verbal reasoning727	.739
13 non-verbal reasoning601	.601
21 rating in English752	.677
22 ratings in mathematics701	.302
15 memory for sentences472	.536
4 common phrases509	.754
5 word pairs575	.729
6 descriptive phrases535	.764
14 memory for digits	—	.375
7 mazes A	—	.222
18 memory for rhythms387	—
16 memory for designs418	—
17 paper folding	—	.263

Evidently, this factor is a reasoning factor with a substantial verbal element, all the verbal tests have high loadings. The loadings corresponding with ratings in English and mathematics are high, as would be expected, but it is of interest that the loading in this factor for the girls' ratings in mathematics is appreciably less high than the others; thus, it may be that reasoning plays less part in mathematical achievement for these girls than for the boys.

The high loadings in memories for rhythms and designs, obtained by the boys only, suggest some differences in strategy between the sexes: possibly boys use verbal description to aid memorizing, whereas girls are more prone to depend on visual, kinaesthetic or auditory perceptions.

The small loadings in mazes A were expected as subjects are likely to use verbal reasoning in the learning of a new technique. The moderate loading in paper folding for girls is of interest since this test involves spatial imagination; for boys the reasoning factor, A, and the spatial factor, C, are highly correlated. It is probably that the experience of boys in technical drawing, crafts with wood and metal, and play or work with machinery may result in an extensive vocabulary enabling them to describe what they perceive in spatial tests, whereas girls who lack experiences and vocabulary depend more on various kinds of imagery.

Test	Factor B	
	Boys	Girls
2 dotted outlines777	.714
5 word pairs526	.336
1 Street Gestalt600	.247
3 mutilated words	—	.672
4 common phrases670	—
6 descriptive phrases657	—
17 paper folding290	—
21 rating in English	—	.382
22 rating in mathematics	—	.300

This factor appears primarily to be a closure factor concerned with speed and strength of closure. Both visual and verbal tests have some high loadings though the factor differs in content for the sexes. In an oblique factor solution Pemberton⁴ distinguished verbal closure, in tests requiring organisation of highly practised symbols, from word fluency; but these factors were correlated. Botzum¹ obtained separate factors of speed of closure and of flexibility of closure with visual materials as well as a factor of verbal fluency, but all three were fairly highly correlated. For the girls, factor B appears close to Pemberton's 'verbal closure'; tests 2 and 3 which require recognition of well-known symbols obtain the only loadings exceeding 0.4. Factor B for the boys is of a wider nature: it includes verbal closure and word fluency, but also closure with visual materials, including symbols, pictures and folded paper. Thus, it corresponds more nearly with Botzum's second order closure factor, though it excludes flexibility of closure.

In this factor also a difference in strategy between the sexes is suggested by the substantial differences in loadings for some tests. Girls obtain high loadings in factor B for the mutilated words tests whereas boys do not. Thurstone⁶ has shown that female students read more quickly than do men; it is probable, therefore, that these girls recognize words more easily than the boys do, so solving this test by rapid closure. For the boys in this investigation this test has high loadings in factor E, flexibility of closure, and in factor C, a spatial factor. A difference in the opposite direction occurs for the Street Gestalt test but it is not so clear cut; this test obtains appreciable loadings also in factor E for boys, but not for girls. Pemberton⁴ observed during individual testing with closure tests that subjects who did not immediately succeed by rapid closure, or a 'synthetic' method proceeded to analyse the figure. Such a contrast between synthetic and analytic solutions has been found by a number of investigators. M. D. Vernon⁷, in surveying these investigations, chiefly with visual materials, suggests that the choice of method may be determined either by test difficulty or by personal bias.

<i>Test</i>	<i>Factor C</i>	
	<i>Boys</i>	<i>Girls</i>
17 paper folding323	.592
10 Gottschaldt II362	.480
15 memory for sentences405	—
14 memory for digits	—	— .519
16 memory for designs673	—
3 mutilated words459	—
6 descriptive phrases204	—
4 common phrases248	—
18 memory for rhythm ..	.277	—
7 mazes A.253	—

Factor C appears to be a spatial factor involving imagination of movement in space and awareness of orientation. For girls the ability tested here is contrasted with memory for digits; that is, a verbal or auditory memory for well-known material is contrasted with visual memories for relatively unfamiliar tasks.

For the boys there are some high loadings also in factor C for memory designs and mutilated words. In discussing visual memory of complex figures Lashley³ points out that they must be broken up and learned piecemeal, so that recall consists, not in a concurrent reproduction of the whole, but in a serial reconstruction of one part after another; he claims that this indicates

that visual memory is a matter of fixation of a direction or a combination of a few directions in visual space. For the boys, factor C is highly correlated with factor A, but it has already been suggested that the training they receive facilitates verbal reasoning with spatial materials.

<i>Test</i>	<i>Factor D</i>	
	<i>Boys</i>	<i>Girls</i>
20 pitch discrimination and memory856	.714
18 memory for rhythms445	.736
19 analysis of chords650	.629
15 memory for sentences	-.334	.382
10 Gottschaldt II324	—
9 Gottschaldt I	—	.300

All very high loadings in factor D are in tests of musical abilities. There is a high loading for girls in the test of memory for sentences, contrasted with a negative loading for boys. Memory for sentences obtains high loadings in factors A and D for girls, but in factors A and C. for boys; it may be that boys imagine the action described in the sentences, visualising it, whereas girls remember by auditory recall. Appreciable loadings in the two Gottschaldt tests are more difficult to account for. Possibly some subjects verbalise their procedure in these tests so depending to some extent on verbal, or auditory, memories. But there is a similarity between these tests and the auditory ones in that a recently played phrase, or chord, or a recently seen pattern must be recognized again. Thus, the factor appears to be not purely auditory but to depend on a more general ability to recall Gestalten after a short interval. It seems not unreasonable that recognition and recall of patterns, whether visual or auditory, may be achieved by some general perceptual ability if we accept the view expressed by Vernon⁸, for example, in a discussion of auditory abilities, that auditory imagery is extremely rare but that laryngeal and other kinaesthetic processes are substituted.

<i>Test</i>	<i>Factor E</i>	
	<i>Boys</i>	<i>Girls</i>
9 Gottschaldt I676	.693
11 Gottschaldt III751	.575
13 A.H.4 non-verbal435	.620
10 Gottschaldt II441	.565
16 memory for designs	—	.538
1 Street Gestalt300	—
3 mutilated words621	—
19 analysis of chords236	—
20 pitch discrimination and memory	—	.238
2 dotted outlines	—	.269

Factor E is closely similar to Thurstone's 'flexibility of closure.' In the case of the Gottschaldt test, this involves recognition of outlines against a distracting background or more complex outlines in which they are included. As in Thurstone's study of visual perception, tests of speed and strength of closure (tests 1-3) also have fairly high loadings in this factor, though these do not appear for both sexes in any of the tests. High loadings in A.H.4 non-verbal reasoning probably appear because a solution must be sought among a number of alternatives which provide distraction as do the complex backgrounds of the Gottschaldt tests.

The auditory test, 19, has a fairly high loading in factor B for girls, but in factor E for boys; and the opposite holds good for test 20. The number of notes in a chord may be swiftly recognized by realising which fingers would be used to play it, for example, or it may be analysed by mentally singing each note. However, the loadings in factor E for the auditory are not substantial and it appears to be in the main a factor of visual perception.

<i>Test</i>	<i>Factor F</i>	
	<i>Boys</i>	<i>Girls</i>
7 Mazes A846	.765
8 Mazes B838	.839
12 A.H.4 verbal355	.230
13 A.H.4 non-verbal432	—
14 memory for digits395	—
9 Gottschaldt I366	—
11 Gottschaldt III290	—
16 memory for designs	—	.430
15 memory for sentences	—	.280
21 rating in English	—	.276
22 rating in mathematics	—	.270
17 paper folding	—	.279

Factor F is defined primarily by the mazes tests. Thus, the claim by their authors that they test an ability not normally measured by intellectual tests appears to be justified; for in this battery we have already obtained factors of verbal reasoning and fluency, a spatial factor and three perceptual factors to which the maze tests make little, if any, contribution. It is true that a test of speed and accuracy in computation was not included but, although it might be expected to obtain a modest loading in this factor, there is no reason to think that it would play a major part in the ability tested here.

Since no other test than the two maze tests has a very high loading, a characteristic of the mazes but common to the other ten tests has been sought. The authors of the mazes suggest that a perceptual scanning mechanism may be employed in solving the mazes and that differences in performance may reflect differences in the size of the perceptual unit an individual can handle and his speed in doing so; in this context scanning implies that each perceptual unit is considered in isolation and then integrated. In factor E consideration of tests with moderate loadings suggests that perceptual span would influence success in them all. Memory for designs which obtains the highest loadings for girls involves scanning as defined here and, if visual memory is employed rather than verbal enumeration of relationships, a greater size of perceptual unit should contribute to success. Similar techniques might be employed in the Gottschaldt tests and in scanning and remembering designs in the A.H.4 non-verbal test. Appreciable loadings in the A.H.4 verbal test, memory for digits and sentences and rating in English suggest that the perceptual span in this factor may be concerned with verbal, or auditory, as well as visual materials. In discussing factor D, it has already been suggested that a common mechanism in visual and auditory skills is provided by kinaesthetic processes.

Until factor F has been obtained in a number of batteries of perceptual tests, it may be difficult to determine its exact nature. In this investigation, the most apt description appears to be 'perceptual span.'

Structure of Perception

<i>Test</i>		<i>Factor G</i>
		<i>Boys</i>
14	memory for digits677
17	paper folding404
15	memory for sentences254
19	analysis of chords306
1	Street Getsalt	-.443
21	rating in English	-.442

<i>Test</i>		<i>Factor G</i>
		<i>Girls</i>
1	Street Gestalt706
21	rating in English	-.213
22	rating in mathematics	-.694
12	A.H.4 verbal	-.350

The nature of this factor which was the last to be extracted is not entirely clear. However, from the largest positive and negative loadings there is a contrast for both sexes between the easy perceptual task of picture completion (test 1) and educational tests. A similar factor was found by Pemberton.⁴ She also showed that a speed of closure factor defined mainly by the Street Gestalt test was associated with a personality unfavourable to school success. Subjects with this personality rated themselves "sociable, quick in reactions, artistic, self-confident, systematic, neat and precise, disliking logical and theoretical problems." For girls rating in mathematics, and for boys rating in English and rote memory, contrast most sharply with this perceptual skill. Since Thurstone⁶ found that ability in the Street Gestalt test did not increase from the third grade upwards the contrast may also be seen as one of difficulty. In general, mathematics is found more difficult by girls and the verbal subjects by boys.

V.—DISCUSSION.

The factors obtained in this analysis include some well validated ones, viz., verbal reasoning, a spatial factor (concerned chiefly with direction or orientation) verbal closure, speed and strength of closure and flexibility of closure. Additional factors obtained are 'perceptual span' and a factor we have described as auditory memory which may have the more general nature of recall of pattern, whether auditory or visual. The factor 'perceptual span' has its greatest loadings in visual tests but has appreciable loadings also in auditory verbal tests. Thus, no factor is concerned solely with visual tests or with auditory tests; but the overlap is not so great as to indicate an inevitable choice of perceptual factors transcending the modalities. However, mental organization itself, as well as experience in learning, may contribute to generalisation of perceptual abilities. Lashley³ suggests that generalisation is one of the primitive, basic functions of organized tissue, for the capacity to perceive proportions and relations among the elements of a stimulus pattern is universal from insects to primates and is retained by rats even after total destruction of the striate cortex. He draws attention also to learning of 'adaptive patterns' which result in instant shifting of a learned reaction from one group of muscles to another, or transposition from visual to motor levels, for example. Since kinaesthetic connections are formed early (and inevitably) in developing visual and auditory perceptions, it seems probable that different aspects of pattern recognition and pattern seeking may be generalised throughout the modalities.

The factors obtained here have high loadings in a greater diversity of tests for boys than for girls. It has been suggested that this may be due to training

which enables boys to verbalise visual experiences. Alternatively, or in addition, a common element in visual and auditory skills may be sought in kinaesthetic cues; where kinaesthetic perceptions have been investigated, for example by Witkin¹¹, boys have been found to excel. It could be that the greater activity of boys, on the average, from infancy contributes to a wider range of kinaesthetic patterns with consequent generalisation to visual and auditory patterns, but this remains speculation until further investigations have been made.

Although we had expected to find that some perceptual factors were more closely associated with mathematical abilities than with ability in English, this is not so. For the boys both sets of ratings gain high loadings in factor A but have no other appreciable loadings. For girls both sets of ratings have high loadings in factor A but also have appreciable loadings in factors B and F. Thus, it appears that success in these subjects for girls depends not only on verbal reasoning, but on verbal closure and on perceptual span. We have suggested, in common with other investigators, that girls depend more on rote learning than boys do.

It is possible that failure to distinguish perceptual factors peculiar to ability in mathematics or English is inevitable here owing to the general nature of the perceptual factors obtained. All five of these factors are concerned with different aspects of pattern recognition or pattern seeking which are perceptual activities fundamental to all learning. Separate factor analyses for the sexes produce substantially the same perceptual factors despite somewhat different experiences of boys and girls and, as we have seen, none of the factors is entirely limited to a single modality. Distinct perceptual factors associated with mathematical or English abilities are more likely to be found in an analysis of test results for an able, homogeneous group of subjects. Possibly one distinction would appear in a contrast suggested by Grey Walter⁹ between skills in pattern-seeking in scientists and mathematicians and in pattern-making among arts specialists, though these capacities are not unrelated.

VI.—SUMMARY OF THE FACTORS OBTAINED.

Factor A: reasoning with visual or verbal materials; including verbal fluency for girls.

Five perceptual factors:

Factor B: (1) *for boys*: speed and strength of closure in recognition of words, objects or phrases in an unorganized or mutilated field.

(2) *for girls*: verbal closure in organizing highly practised symbols.

Factor C: a spatial factor involving awareness of direction or orientation.

Factor D: recognition and recall of Gestalten chiefly in auditory tests.

Factor E: flexibility of closure in recognition of a pattern amid distracting and confusing detail in the field, mainly with visual materials.

Factor F: Perceptual span in visual or verbal tests, whereby span is meant the size of perceptual units which are 'scanned' and which may be integrated.

Factor G: contrasting success in an easy perceptual task of picture completion with attainment in school subjects.

ACKNOWLEDGEMENTS.—The author is grateful to Dr. A. Elithorn and to Mrs. Kerr for the use of test materials, discussion of difficult points, and for patience in waiting for the results. Thanks are also due to colleagues who helped with ideas and materials: Miss E. Law and Dr. A. Wilkinson for suggesting phrases for tests 4–6, Dr. J. Brocklehurst for advice on tests of auditory abilities, Mr. C. J. Phillips for aid in making tapes, Dr. H. Maddox for providing references and help with apparatus, Mrs. J. Goode for aid in teleprinting and Mr. J. Hallworth for providing computing programmes.

VII.—REFERENCES.

- ¹ BOTZUM, W. A. (1951). A factorial study of the reasoning and closure factors. *Psychometrika*, Vol. 16, pp. 361–386.
- ² ELITHORN, ALICK, KERR, MYFANWY, and MOTT, JOY (1960). A group version of a perceptual maze test. *Brit. J. Psychol.*, 51, 1, 19–26.
- ³ LASHLEY, K. S. (1942). The problem of cerebral organisation in vision. *Biol. Symposia*, 7, 301–3321.
- ⁴ PEMBERTON, CAROL (1948). The closure factors related to intelligence. *Psychometrika* (March), 13, 1.
- ⁵ SLATER, PATRICK (1964). Specific factors and minor components. *Brit. J. Psychol.*, Vol. 55, Part 3 (Aug.).
- ⁶ THURSTONE, L. L. (1944). *A Factorial Study of Perception*. University of Chicago Press.
- ⁷ VERNON, M. D. (1962). *The Psychology of Perception*. Pelican.
- ⁸ VERNON, P. E. (1935). Auditory perception: II.—the evolutionary approach. *Brit. J. Psychol.*, 25, 3, 265–283.
- ⁹ WALTER, W. GREY (1961). *The Living Brain*. Pelican.
- ¹⁰ WITHRINGTON, J. (1936). *An Analysis of Factors Entering into Geometrical Ability*. M.A., London.
- ¹¹ WITKIN, H. A., et al. (1954). *Personality Through Perception*. New York: Harper.

(Manuscript received 22nd November, 1964.)

TABLE I

		INTER TEST CORRELATIONS.																					
		BOYS											GIRLS										
Test		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	053																						
2	319	150																					
3	241	165	167																				
4	440	509	167	188																			
5	229	472	188	666	717																		
6	343	570	290	769	262	301	257																
7	186	264	220	262	262	280	200	701															
8	225	269	220	189	280	200	229	268	396														
9	294	236	281	121	231	229	229	268	396	254													
10	141	304	395	408	399	394	346	264	354	254	374												
11	179	115	318	141	102	189	229	364	455	374	222	212	015	061	247	127	030	163	206	-056	-114		
12	223	255	300	536	559	576	422	299	342	355	444	416	713	682	370	401	150	117	186	009	179	601	515
13	394	364	398	521	525	538	497	443	486	444	289	344	344	305	255	291	158	108	088	102	284	279	241
14	066	288	080	336	244	246	325	376	148	348	289	344	344	305	255	291	158	108	088	102	284	279	241
15	063	102	197	380	294	365	250	204	159	177	147	365	292	274	260	145	029	-020	-072	158	297	241	
16	059	066	340	385	348	357	337	163	079	450	237	345	436	232	368	183	232	319	301	093	205	459	090
17	176	315	421	346	333	377	200	282	265	400	251	334	339	346	240	354	165	129	189	192	298	179	191
18	220	301	294	446	345	389	270	177	131	354	210	500	497	383	338	293	165	129	189	192	298	179	191
19	102	230	144	338	383	338	338	097	144	183	304	213	164	312	258	129	189	192	298	179	191	191	
20	096	242	034	274	286	261	261	226	060	242	020	244	181	080	-139	224	087	395	491	289	043	-032	-032
21	-109	159	132	486	546	504	147	060	026	286	-017	427	350	313	319	345	305	288	419	216	148	096	560
22	-101	186	098	289	339	289	034	164	192	222	080	425	470	142	085	090	091	212	243	290	546		

Correlations in thousands.

Boys' mean correlation : .285, omitting ratings (21, 22) .296 (73 boys).

Girls' mean correlation : .172 .165 (72 girls).

TABLE 2

FACTOR LOADINGS OBTAINED BY VARIMAX ROTATION.

Test	Boys							Girls						
	I	II	III	IV	V	VI	VII	I	II	III	IV	V	VI	VII
1	.600	.319	-.045	.009	.175	-.060	-.443	.175	-.083	-.245	.233	-.105	.706	.214
2	.763	.148	-.156	.177	.176	-.035	.169	.016	-.269	-.093	.728	.239	-.093	-.078
3	.113	.623	.481	.047	-.041	-.022	-.147	.075	-.156	.156	.670	-.182	.121	.104
4	.783	.000	.322	.144	.086	.232	.107	.777	.026	-.045	.010	-.072	.188	.002
5	.638	.025	.240	.171	.122	.391	.157	.799	.030	.078	.187	.067	-.138	.168
6	.772	.109	.298	.106	.038	.289	.091	.808	.069	.100	.021	.102	.054	.052
7	.119	.077	.284	.132	.842	.009	.033	.261	.104	.148	.089	.143	-.038	.765
8	.098	.285	-.017	.087	.814	.065	.150	.185	.092	-.044	.030	.020	-.015	.839
9	.130	.682	-.149	-.070	.300	.284	-.057	.057	.691	.303	-.127	-.196	.015	-.166
10	.218	.448	.393	.324	.088	.045	.244	.106	.557	-.089	.000	.480	-.128	.200
11	-.034	.741	.078	.053	.217	.032	.149	-.078	.586	.123	.129	.178	.158	.198
12	.346	.152	.393	.024	.341	.564	-.039	.738	.064	.051	-.172	-.206	-.350	.230
13	.347	.426	.303	.099	.392	.493	-.079	.629	.560	-.239	-.206	-.078	-.177	.020
14	.192	.150	.102	.076	.382	.060	.677	.401	.144	-.050	.028	-.519	.019	.153
15	.218	-.005	.555	-.334	.200	.1193	.254	.557	.029	.329	-.154	.122	.193	.280
16	.079	.146	.792	.171	.104	.049	.094	.129	.528	.048	.279	-.151	.091	.430
17	.348	.468	.284	-.026	-.008	-.034	.404	.281	.064	.075	.014	.592	.142	.269
18	.277	.099	.405	.445	.200	.172	.021	.170	-.028	.722	-.152	.055	.026	.109
19	.198	.207	.048	.650	-.085	.205	.306	.024	.103	.631	.206	.045	-.147	-.158
20	.131	-.072	.047	.856	.187	.126	-.087	.071	.232	.713	.128	.200	-.002	.130
21	.228	-.104	.321	.173	-.115	.640	.442	.732	-.211	.069	.242	.018	-.213	.276
22	.100	.109	-.074	.202	-.002	.848	.021	.352	.003	-.070	.235	-.216	-.694	.270

% Variance extracted.

15.03 10.78 10.55 8.25 9.83 10.02 6.36 19.34 9.06 8.23 6.75 5.64 6.35 9.57

TABLE 3

THE FINAL FACTOR LOADINGS.

Test	Boys							Girls						
	A	B	C	D	E	F	G	A	B	C	D	E	F	G
1	.023	.600	.072	.009	.300	.206	-.443	.140	.247	-.105	-.227	.100	.214	.706
2	.041	.777	-.027	.177	.135	.190	.169	-.144	.714	.239	-.094	.269	-.078	-.093
3	.193	.017	.459	.047	.621	.021	-.147	-.059	.672	-.182	.164	-.148	.104	.121
4	.509	.670	.248	.144	.023	.086	.107	.754	.161	-.072	.032	.103	.002	.188
5	.575	.526	.085	.171	.058	.124	.157	.729	.336	.067	.156	.109	.168	-.138
6	.535	.657	.204	.106	.141	.049	.091	.764	.174	.102	.179	.149	.052	.054
7	.166	.059	.253	.132	.006	.846	.033	.222	.110	.143	.172	.130	.765	-.038
8	.050	.091	-.026	.087	.209	.838	.150	.181	.006	.020	-.026	.110	.839	-.015
9	.143	.120	-.222	-.072	.676	.366	-.057	-.016	-.133	.196	.300	.693	-.166	.015
10	.230	.130	.362	.324	.441	.132	.244	.056	.021	.480	-.084	.565	.200	-.128
11	-.004	-.052	.086	.053	.715	.290	.149	-.170	.098	.178	.109	.575	.198	.158
12	.727	.190	.103	.024	.176	.355	-.039	.739	-.027	-.206	.123	.137	.230	-.350
13	.601	.218	.074	.099	.435	.432	-.079	.601	-.090	-.078	.018	.620	.020	-.177
14	.129	.160	.085	.076	.119	.395	.677	.375	.104	-.519	-.011	.183	.153	.019
15	.472	.080	.405	-.334	-.004	.199	.254	.536	-.050	.122	.382	.084	.280	.193
16	.418	-.083	.673	.171	.138	.118	.094	.015	.291	-.151	.055	.538	.430	.091
17	.135	.290	.323	-.026	.464	.039	.404	.263	.066	.592	.102	.092	.269	.142
18	.387	.169	.277	.445	.098	.209	.021	.127	-.130	.055	.736	-.011	.109	-.026
19	.220	.158	.027	.650	.236	.064	.306	.089	.192	.045	.629	.105	-.157	-.147
20	.164	.103	-.012	.856	-.076	.179	-.087	-.048	.121	.200	.714	.238	.130	-.002
21	.752	.080	-.019	.173	-.026	.125	.442	.677	.382	.018	.143	-.137	.276	-.213
22	.701	.007	-.257	.202	.193	.009	.021	.302	.300	-.216	-.035	.038	.270	-.694

% Variance extracted

16.5 11.1 6.7 8.2 10.5 10.8 6.4 17.4 7.5 5.7 8.8 9.7 9.6 6.3

A STUDY OF SOME CONTRIBUTIONS OF THE MORRISBY DIFFERENTIAL TEST BATTERY TO VOCATIONAL SELECTION

By M. K. L. EL-SHARKAWY AND D. M. LEE

(University of London Institute of Education)

SUMMARY. This paper reports an investigation concerned with gathering information about the Morrisby Differential Test Battery in relation to various claims for its usefulness in vocational selection. Along with a pre-constructed attitude scale, the complete test battery was administered to 154 entrants and 163 apprentices of three years' standing in several large firms within the British Iron and Steel Industry. In particular, the scores describing the Morrisby Modal Profiles were compared with Training Officers' ratings of similar characteristics. In the present paper the main analyses are from scores obtained by the older apprentices; some of the data on reliability coefficients were gained by retesting a group of entrants. It is the intention of the authors to follow up the entrants at the end of three years to obtain further information about the predictive validity of this test battery and to report this in a later paper.

The individual reliability coefficients of the twelve tests comprising this battery are shown to vary widely. In view of the relatively small number of testees, these results can only be regarded as indicative and further studies are needed. The test manual lacks information about the validity of the battery, though extensive claims are made for the so-called 'modal profiles' in estimating the intellectual 'personality' of the individuals concerned. Against the two criteria here employed, the 'ability' tests are shown to have more validity than the 'personality' profiles. The 'skill' tests show very low validity coefficients.

Morrisby stresses the differential nature of his battery and suggests that a basic factorial structure is neither present nor intended. A preliminary factor analysis of scores obtained from this investigation showed four main factors of a type generally found in the analysis of large test batteries. A profile chart could be constructed on the basis of these factors which could be of use in vocational selection.

I.—INTRODUCTION.

EDUCATION is concerned with training. At the higher levels, training for any specific career is suited to some individuals rather than others; this kind of training depends on the prior selection of suitable persons. Thus, selection is one important facet of training, and good selection procedures considerably facilitate the training process itself. Such a process may be said to include any activity or series of experiences designed to improve the individual's performance in a particular task, skill or situation.

In large industries, the spread of apprenticeship is wide and many young workers seek entry. A selection procedure may exist for each type of apprenticeship, in which case boys may try for several before they are accepted. This system of selection may be wasteful in time, manpower and money. More often one selection procedure must identify groups of boys for each apprenticeship in the range offered. Such a procedure makes extensive demands on psychological resources of a type which have only recently begun to be considered. That such procedures are essential is apparent far beyond the apprenticeship field. Similar situations are found in all aspects of employment or higher education. Individual problems may vary but the need for a common procedure remains.

II.—OUTLINE OF PREVIOUS RESEARCH.

Since the early days of test construction there has been a growing interest in the field of occupational selection. Historically, this proceeded by job analysis first, but initially the tests were sensory rather than intellectual in nature. With the growth of work on intelligence tests, however, studies began to be made concerning the relations between intelligence and occupational levels. Kornhauser¹ and Yoakum² describe some of these studies. As a result, there were a number of attempts to classify vocations, according to the apparent degree of intelligence required.

The development of factor analysis techniques gave a powerful impetus to the construction of tests for vocational selection. In the general field of mental testing, efforts were made to isolate separate group factors. A considerable volume of research suggested that, far from being a single factor, intelligence was more likely to be the result of a combination of verbal, numerical, practical and spatial visualisation factors. Seashore,³ Stenquist,⁴ and Cox⁵, attempted to establish mechanical ability as a sub-combination of these group factors. Farmer⁶, Earle⁷, and Gaw⁸ worked on tests of manual dexterity and motor ability. El-Koussy⁹ and Thurstone¹⁰ attempted an analysis of the factor of spatial visualisation. All these factors were seen to contribute to the assessment of mechanical ability, and some level of success in the corresponding tests appeared to be necessary for entry into certain types of occupation. There were similar investigations into potential tests for suitability for clerical work.

The early work of Bingham and Davis¹¹ had shown the importance of certain personality factors in the vocational selection field. Some years later, the work of such investigators as Fryer¹² and Strong¹³ supported the importance of such factors, and interest tests were developed. These were seen to make a useful contribution, though difficulties were speedily recognised in relation to such matters as the faking of answers or questions of validation for selection purposes. Maier¹⁴ contributed similar investigations on attitude scales, but as these were usually measured after entering a particular occupation, the contribution of this work for actual vocational selection purposes was somewhat lessened.

In 1928, Hull¹⁵ had suggested the need for a comprehensive battery of tests and test weights for the broad occupational field. In this respect, he looked forward to a development which only seriously began to take shape after 1939, when the growth in special aptitude tests began to suggest the idea of custom-built batteries of tests for specific occupational fields. In the last fifteen years the trend has moved away from this towards the creation of differential aptitude batteries to service a range of occupational fields, this latter development being based on the dual realisation that a complexity of abilities is required for success in any particular occupation and that there is overlap between the abilities brought into play in several differing vocational fields. The early differential aptitude test batteries originally utilised tests of intellectual abilities, but later ones claim additional personality interpretations.

Since the publication, in 1941, of the Thurstone tests of Primary Mental Abilities, at least six¹⁶ major differential aptitude batteries have been tried. The most well-known are probably the Guilford—Zimmerman Aptitude Survey and the Flanagan Aptitude Classification Tests. For any one of these series of tests only limited evidence of differential validity is available for any type of vocational selection. The most recent battery to be added to the above list is that of Morrisby¹⁷ known as the Differential Test Battery (D.T.B.). This

is a mixed battery of tests for which very considerable occupational claims are made. It is the purpose of this paper to review some evidence on the reliability and validity of this battery and its claims.

III.—PLAN OF THE INVESTIGATION.

(i) *Subjects and Materials.*

The subjects used in this investigation comprised 154 new entrants (aged 15 and 16 years) and 163 apprentices aged 19 years. For the purposes of the present paper only a small group of the new entrants who were retested for test reliability calculations have been used; the main interest attaches to the 19-year-old groups. These young men were drawn from seven firms, all within the iron and steel industry, all located within one geographical area, and all carrying out similar apprenticeship schemes leading to comparable types of employment. The firms all volunteered to take part in the experiment and covered both small and large undertakings.

The complete Morrisby Differential Test Battery was used in precisely the manner laid out in the manual. This battery represents the multi-occupational approach and, therefore, apparently makes it possible to explore a variety of occupational possibilities. It is devised as an instrument for the assessment of human ability structures for the prediction of the performance of persons in occupations or courses, and for determining the relative suitability of various occupations and training courses for an individual person. It is composed of twelve tests as follows:

(1) The Compound Series Test	(C.S.T.)
(2) General Ability Test—Verbal	(GAT-V)
(3) General Ability Test—Numerical	(GAT-N)
(4) General Ability Test—Perceptual	(GAT-P)
(5) Shapes Test	(Shapes)
(6) Mechanical Ability Test	(Mech.)
(7) Speed Test 1, Number and Name Checking	(No. and Na.)
(8) Speed Test 2, Perseveration	(Pers.)
(9) Speed Test 3, Word Fluency	(Wd. Fl.)
(10) Speed Test 4, Ideational Fluency	(Id. Fl.)
(11) Speed Test 5, Motor Speed	(Mot. Sp.)
(12) Speed Test 6, Motor Skill	(Mot. Sk.)

The first four tests are of the usual intelligence type. Test 1 is a basic test, and the other three are said to present a differential assessment of the three main areas of mental ability since they cover comparably the three main media of verbal, numerical and diagrammatic material. Test 5 is similar to the tests of spatial ability produced by El-Koussy and Thurstone. Test 6 is said to be a test of 'mechanical insight.' This pair of tests is said to be valuable both as positive predictors in the practical and manual fields as well as indicators of converse relationships with tests of other kinds of capacities. The first six tests are all tests of 'intellectual capacity.' The last six tests are speed tests contained in one booklet; these are said to be tests of 'intellectual personality,' or the manner in which the testee undertakes an intellectual performance. Test 7 is one common in batteries of tests for clerical aptitude. Tests 8, 9, 10 are similar to some found in intelligence tests. Test 11 is a 'pure' speed test, and Test

12 is a speed test requiring some measure of intellectual control. It must be emphasised, however, that while there are similarities in the content of these speed tests and other well-known tests, in the Morrisby Differential Test Battery they are intended for use in a particular way. It is essential that the instructions in the manual be followed completely in the administration of these six tests.

Morrisby has stated that the Differential Test Battery is not factorial in nature. A system of scoring has been evolved in which two aspects are important: (i) the level of individual score, in each case formulated as a scale score comparable from test to test; (ii) the difference between these scale scores for certain pairs or groups of tests.

Thus, interest centres round the particular interpretations to be accorded to certain patterns or profiles of scores. Notes on typical score profile patterns suggested by the author have been derived from the observation of a large number of cases in which the observed profile pattern could be reconciled with such other attributes as known characteristics and personal achievement. The following are the typical score profile patterns suggested to users of the Morrisby battery, but the proportions shown are not intended to be absolute:

- (1) A low score on C.S.T. indicates the academic, book-learning type. A high score indicates the type of individual more likely to succeed with subjects requiring 'intellectual power' (e.g., Mathematics) and also more likely to succeed with advanced study. It is suggested, however, that this score is not to be taken as definitive, but rather in conjunction with other profiles mentioned below.
- (2) Profiles for the three general ability tests may be viewed as follows:

	Arts	Science	Commerce	Social Science
GAT-V	XXXXXXXX XXXXXXXX	XXXX XXXX	XXXXX XXXXX	XXXXXX XXXXXX
GAT-N	XXXXXX XXXXXX	XXXXXX XXXXXX	XXXXXXX XXXXXXX	XXXX XXXX
GAT-P	XXXX XXXX	XXXXXXXX XXXXXXXX	XXXXX XXXXX	XXXXXXX XXXXXXX

- (3) Dual profiles for the shapes and mechanical ability tests:

	Craft	Mechanic
Shapes	XXXXXX XXXXXX	XXXX XXXX
Mech.	XXXX XXXX	XXXXXX XXXXXX

- (4) Modal Profiles for differing types of 'intellectual personality' as determined by Tests 7-10 are suggested as follows :

TYPE A AND VARIATIONS.

	Type A Normal	Variations		
		Scatter-brained	Inspid	Salesman
No. and Na.	XXXXXX XXXXXX	XXXXXXXXXX XXXXXXXXXX	XXXXXX XXXXXX	XXXX XXXX
Pers.	XXXX XXXX	X X	XXXX XXXX	XXXX XXXX
Wd. Fl.	XXXXXX XXXXXX	XXXXXXXXXX XXXXXXXXXX	XXXXXX XXXXXX	XXXXXX XXXXXX
Id. Fl.	XXXXXX XXXXXX	XXXXXXXXXX XXXXXXXXXX	XXXX XXXX	XXXXXX XXXXXX

TYPE B AND VARIATIONS.

	Type B Normal	Variations		
		Rigid	Tear-Away	Executive
No. and Na.	XXXX XXXX	XX XX	XXXX XXXX	XXXXXXXX XXXXXXXX
Pers.	XXXXXX XXXXXX	XXXXXXXXXX XXXXXXXXXX	XXXXXXXXXX XXXXXXXXXX	XXXXXXXX XXXXXXXX
Wd. Fl.	XXXX XXXX	XX XX	XXXX XXXX	XXXXXXXX XXXXXXXX
Id. Fl.	XXXX XXXX	XX XX	XXXXXXXXXX XXXXXXXXXX	XXXXXXXX XXXXXXXX

Six personality characteristics¹⁸ are attached to each variation in each case.

- (5) The two final speed tests are said to give dexterity profiles with interpretation thus :

	Speed	Skill
Mot. Sp.	XXXXXXXX XXXXXXXX	XXXX XXXX
Mot. Sk.	XXXX XXXX	XXXXXXXX XXXXXXXX

It is suggested that the Speed Profile is indicative of routine operatives, whereas the other points to occupations requiring manual skill.

Few details of reliability or validity seem to be available for the above tests. Some split-half and some re-test reliabilities are quoted, but these are not given for all tests. The validity of the profiles would appear to rest entirely on the subjective interpretation of observed cases, and only general statements are made. Thus, the information on the reliability and validity of the battery is by no means complete. The most informative section of the manual deals with the actual administration of the tests and this is straightforward and clear.

In order to gain additional information on the validity of the test battery, further criteria were needed. The profiles associated with the first six tests seemed not unreasonable in the light of the present results in the field of ability testing. In any case, a factorial analysis of the data and a simple comparison with Training Officers' Ratings could throw further light on these. As criteria for the particular assessment of the modal profiles it was decided to construct a formal attitude scale and to obtain Training Officers' Ratings for a range of related personality characteristics. The Employees' Attitude-Dimensions Scale (E.A.D.S.)¹⁹ was constructed especially for this investigation, using a modified form of the Likert technique. It contains twelve items on each of eight dimensions:

- | | |
|----------------------------------|--------------------------------------|
| (1) Job. | (5) Internal Society. |
| (2) Management and Organisation. | (6) Participation and Responsibility |
| (3) Working Conditions. | (7) Supervision. |
| (4) Incentives. | (8) External Society. |

The content was such that some correspondence with the modal profiles could be established. No time-limit was set, but the scale was usually completed in 20-25 minutes. Using Rulon's formula, the reliabilities of the eight individual sections varied between .66 and .94.

All the 19-year-old subjects were of three years' standing in their respective firms and hence were well-known to the Training Officers concerned. These Training Officers were asked to rate each of the subjects in their charge on eight ability characteristics and fourteen personality characteristics. The eight ability characteristics related individually to Tests 1-6 and 11, 12. The fourteen personality characteristics included all those listed by Morrisby as underlying the so-called modal profiles, thereby supplying an alternative frame of reference for the assessment of these. A check-sheet of the customary type was provided for the recording of these ratings.

(ii) *Experimental Procedures.*

The Differential Test Battery was administered to both new entrants and apprentices of three-years' standing within three weeks in September/October, 1962. There were seventeen sessions at the rate of one or two per day, ranging over seven firms. In each case, the Morrisby Test was completed in one session lasting from 3 hours 15 minutes to 3 hours 30 minutes. The Training Officers each informed their own apprentices of the purpose of the investigation, and of the date, time and place of meeting. Each firm loaned one or two staff who helped very considerably with the arrangement of rooms and test materials and with many matters of general organisation necessary for the smooth running of the experiment. The actual administration of the tests was carried out by two people who had previously taken the Morrisby week-end training course.

The Employees' Attitude-Dimensions Scale was administered to the 19-year-old group after the completion of the Differential Test Battery. No member of the staff of any firm was present. Subjects were asked to complete

information concerning their names, type of work, course of training, etc., so that the scales could be collated with other data. This information was sought after the completion of the attitude scale. Lastly, a meeting was held with the Training Officers to discuss the rating schedule and problems relating to its completion. Each of them then completed a rating schedule for each apprentice in their charge during the weeks following the administration of the test battery and attitude scale. For this purpose they were asked to take account of all evidence at their disposal.

IV.—STATEMENT OF RESULTS.

(i) *Preliminary Analysis of the Measurement Data.*

The sample of 19-year-old apprentices used in this investigation was drawn from seven companies within the iron and steel industry. Taking the largest group as criterion, five of the remaining groups did not differ significantly at the 5 per cent. level for either means or standard deviations. One group showed a highly significant departure in respect of the mean; this company was known to have selected its superior apprentices. The full range of information was not available for another company. These groups have been discarded for the statistical analysis, leaving a sample of 124 apprentices for this purpose. This latter sample may be taken to represent a randomly drawn sample of apprentices throughout the industry at large.

An examination of individual distributions showed that all the individual tests of the Differential Test Battery were well suited to the sample, whose scores gave rise to distributions of the normal type. The manual of the test battery gives split-half reliabilities for some of the tests, and one re-test reliability for the Compound Series Test. In the present investigation, re-test reliability coefficients were obtained from retesting a small group of eighteen new entrants after an interval of ten days. This is a small group and the results can only be taken as exploratory. Table 1 shows the available reliability coefficients from the manual and from the present investigation.

The reliability coefficients quoted in Table 1 from this investigation were obtained by the application of Rulon's formula, this seeming the most appropriate formula in this case. The numerical coefficients appear reasonably satisfactory, more especially since re-test reliabilities are in general lower than their split-half counterparts, and since the small numbers would also tend to lower the coefficients.

The Employees' Attitude-Dimensions Scale was examined in the light of the individual distributions of scores for each of the eight sections. These distributions were all satisfactory for the main sample and indicated a healthy variation in attitude in each case. The scores for all eight sections were inter-correlated in order to examine the pattern or cohesion within the scale. Three major groupings were shown:

A. Job participation Responsibility External Society	B. Supervision Internal Society Management and Organisation	C. Incentives Work Conditions
--	--	----------------------------------

This grouping supports the logic behind the construction of attitude scale. The discriminating quality of the attitude scale was further checked by examining the relative means for each section in the top-third (high attitude) and lower-third (poor attitude) groups. In each case, the difference between means was significant at the 5 per cent. level or above; this gave sufficient indication of the discriminatory quality of the individual sections.

TABLE 1

RELIABILITY COEFFICIENTS FOR THE MORRISBY DIFFERENTIAL TEST BATTERY.

Test	Reliability Coefficients from the Manual		Reliability Coefficients from this investigation
	split-half	re-test	re-test
C.S.T. ..	.984	.930	.67
GAT-V ..	—	—	.79
GAT-N ..	—	—	.79
GAT-P ..	—	—	.91
Shapes ..	.962	—	.88
Mech. ..	.897	—	.90
No. & Na.	—	—	.86
Pers.	—	—	.88
Wd. Fl. ..	—	—	.94
Id. Fl. ..	—	—	.87
Mot. Sp. .	—	—	.63
Mot. Sk.	—	—	.73

The twenty-two characteristics comprising the Training Officers' Rating Scale were also examined for suitability in terms of the individual distributions. These distributions were all normal in type, this possibly only reflecting adherence to the initial instructions for rating. The scale appears to have operated successfully in this context. Since this scale is to be used with the Employees' Attitude-Dimensions Scale to provide criteria for assessing the validity of the Differential Test Battery, it seemed wise to try to assess the relationship between the two scales themselves. Ferguson's version of the Mann-Whitney test showed that there is some relationship between the general level of attitudes as measured by the Employees' Attitude-Dimensions Scale and some of the personality characteristics rated by the Training Officers. This relationship, however, is limited to four main characteristics, and there is thus sufficient independence for the two scales to provide separate estimates of the validity of the Differential Test Battery.

(ii) *Validity of the Differential Test Battery.*

The profiles claimed by Morrisby, in relation to the six ability tests and the two skill tests in the battery seem reasonably well-based in the logic of those tests as already supported by a considerable volume of research, and further light may be thrown on these by a comparison with the Training Officers' Ratings and by the results of the factorial analysis described later. The modal profiles based on the first four speed tests appear rather less well supported, and require careful investigation with all possible measures. The results of the Employees' Attitude-Dimensions Scale and the Training Officers' Ratings present criteria for the further evaluation of the validity of the three main aspects of the Morrisby Differential Test Battery. These results are now examined in turn.

Three methods were used for checking the modal profiles against the results of the Employees' Attitude-Dimensions Scale. Using the product-moment correlation coefficient, the scores on each of the four speed tests were correlated in turn with the scores on each dimension of the Employees' Attitude-Dimensions Scale. Out of thirty-two correlation coefficients, five were just significant at the 5 per cent. level and two of these were negative. The significant correlation coefficients are shown in Table 2. It will be seen that nothing can be gained from this table in regard to the modal profiles.

For a second check on the modal profiles against the Employees' Attitude-Dimensions Scale, the subjects were ranked in order of total scores on the

TABLE 2

CORRELATIONS BETWEEN THE FOUR MODAL PROFILE TESTS AND THE EIGHT DIMENSIONS OF THE EMPLOYEES' ATTITUDE-DIMENSIONS SCALE.

Tests	No. and Na.	Pers.	Wd. Fl.	Id. Fl.
Dimensions				
Job	—	—	—	—
Man. and Org.19	— .18	—	—
Work Cond.	—	—	—	—
Incentives	—	— .17	—	—
Int. Soc.	—	—	—	—
Part. and Resp.21	—	—	—
Supervision	—	—	—	.18
Ext. Soc.	—	—	—	—

attitude scale. Those in the top third were then taken as a group with high or positive attitudes to their chosen occupation, and those in the lower third were taken as a group with low attitudes. For each of the modal profile tests, the significance between the means of the high and low attitude groups were calculated by means of the t-test technique. Two tests showed a difference between means which was significant at the 5 per cent. level; these were Number and Name Checking (No. and Na.) and Ideational Fluency (Id. Fl.). This particular pattern of significance is not reflected in any of the profiles described by Morrisby.

A method described by Ferguson²⁰ for estimating the difference between two proportions was employed for a third check on the modal profiles using the attitude scale. The highest and lowest groups on the Employees' Attitude-Dimensions Scale were again taken to represent groups showing high and low attitudes to their everyday work. Profiles of all subjects in the high and low attitude groups were examined for their relevance to the eight profiles described by Morrisby. The proportions in the high and low attitude groups were noted for each of the eight profiles, and the significance between the proportions judged by the Ferguson test. The results are shown in Table 3.

TABLE 3

Z VALUES OF THE DIFFERENCES BETWEEN THE PROPORTIONS OF MODAL PROFILES FOUND AMONGST HIGH AND LOW ATTITUDE GROUPS ON THE EADS SCALE.

	Proportions "A" Profiles				Proportions "B" Profiles			
	Normal	Scatter-brained	Insipid	Sales-man	Normal	Rigid	Tear-away	Executive
H	.37	.24	.20	.02	.05	.00	.10	.02
L	.20	.07	.24	.05	.20	.02	.05	.17
Z	2.01	2.66	.44	.60	2.67	1.01	.75	2.25
	Sig. .05	Sig. .01	—	—	Sig. 0.1	—	—	Sig. .05

It should be noted that the two significances for the A-type profile are in the opposite direction to those for the B-type profile. It should further be noted that the A-type profiles on the whole tend to occur more often than the B-types.

The validity of the Differential Test Battery was also checked against the Training Officers' Ratings for the eight corresponding abilities and for the fourteen personality characteristics which encompassed all those suggested by Morrisby in connection with the interpretation of the modal profiles. For ease of discussion three groupings of the individual tests in the battery may be used: the ability tests (Nos. 1-6), the speed tests (Nos. 7-10), and the skill tests (Nos. 11, 12). This grouping corresponds to the profiles obtainable from the use of the battery as a whole. The product-moment correlation coefficients between the individual tests of the battery and the various Training Officers' Ratings are given in Tables 4-6 below. Only coefficients significant at the 5 per cent. level are shown.

TABLE 4

CORRELATION COEFFICIENTS BETWEEN THE SIX "ABILITY" TESTS OF THE DIFFERENTIAL TEST BATTERY AND THE TRAINING OFFICERS' RATINGS ON THESE ABILITIES.

Co.	1	2	3	4	5	6	Total Sample
Test							
C.S.T.	—	.46	—	—	—	—	.32
GAT-V	—	—	.59	.44	.44	.66	.34
GAT-N	—	—	—	—	—	—	—
GAT-P	—	.39	.54	.46	—	.55	.41
Shapes	—	.39	.41	—	—	—	.20
Mech.	—	—	.61	—	—	.61	.37

In the above table it will be seen that the number test GAT-N shows no significant correlation with the corresponding ratings. The verbal test GAT-V and the perceptual test GAT-P show the most consistent pattern of relationships in this field.

TABLE 5

CORRELATION COEFFICIENTS BETWEEN THE TWO "SKILL" TESTS OF THE DIFFERENTIAL TEST BATTERY AND THE TRAINING OFFICERS' RATINGS ON THESE ABILITIES.

Co.	1	2	3	4	5	6	Total Sample
Test							
Mot. Sp.	—	.34	—	—	—	—	—
Mot. Sk.	—	—	—	—	—	.50	—

From tests 11 and 12 only two significant correlation coefficients are present, and no consistent pattern is shown (Table 5). It will be observed that there are no patterns of significant correlation coefficients across the modal profile tests. Of the individual speed tests only the Number and Name Checking shows several significant correlation coefficients. The tests of word fluency and ideational fluency show no significant correlation coefficients with the personality characteristics.

TABLE 6

CORRELATION COEFFICIENTS BETWEEN THE MODAL PROFILE (SPEED) TESTS OF THE DIFFERENTIAL TEST BATTERY AND THE TRAINING OFFICERS' RATINGS OF THE FOURTEEN PERSONALITY CHARACTERISTICS.

Personality Characteristics	Modal Profile Tests.			
	No. and Na.	Pers.	Wd. Fl.	Id. Fl.
1. Free and easy in general approach	.24	— .19	—	—
2. Operates in terms of rules22	—	—	—
3. Imaginative30	—	—	—
4. Co-operative19	—	—	—
5. Reliable	—	—	—	—
6. Willing to accept criticism	—	—	—	—
7. Flexibility of mind25	—	—	—
8. Steady	—	—	—	—
9. Willing to accept responsibility ..	—	—	—	—
10. Amenable to authoritarian discipline	—	—	—	—
11. Rigid	—	—	—	—
12. Ambitious18	—	—	—
13. Tactful	—	—	—	—
14. Active under most circumstances	—	—	—	—

For each of the modal profiles, the combined correlation coefficients between each of the speed tests and the related group of personality characteristics as ascribed by Morrisby were calculated with the use of z coefficients..

TABLE 7

COMBINED CORRELATION COEFFICIENTS BETWEEN THE FOUR SPEED TESTS AND THE GROUPS OF PERSONALITY CHARACTERISTICS AS ASCRIBED BY MORRISBY AND JUDGED BY THE TRAINING OFFICERS.

	Type A Normal		Type B Normal	
	Profile	Comb. Corr.	Profile	Comb. Corr.
No. and Na.	xxxxxx xxxxxx	0.21	xxxx xxxx	—
Pers.	xxxx xxxx	—	xxxxxx xxxxxx	—
Wd. Fl.	xxxxxx xxxxxx	—	xxxx xxxx	—
Id. Fl.	xxxxxx xxxxxx	—	xxxx xxxx	—

Table 7 shows two examples of the apparent relationships between the modal profiles on the one hand and the ratings of the same grouping of personality characteristics by the Training Officers. There is only one correlation coefficient which is significant at the 5 per cent. level. The balance of tests suggested in the modal profiles is not borne out by the comparable evidence obtained from the Training Officers' Ratings. Thus, the personality characteristics would seem to bear little relation to the speed tests underlying the modal profiles.

In a second approach to the examination of the modal profiles by means of the Training Officers' Ratings for the fourteen personality characteristics, the subjects were grouped under six profiles according to where they seemed most appropriately placed by their relevant test scores. From the information gained from the Training Officers' Ratings it could then be seen whether any particular personality characteristics stood out for any one profile. Table 8 shows which characteristics seemed markedly positive or negative for which profiles, x positive, -negative.

It will be seen from Table 8 that certain positive personality characteristics are important in Morrisby's 'A' profiles and certain negative characteristics in the 'B' group. Though the pattern in Table 8 does not agree with the individual profile patterns suggested by Morrisby, there is a reasonable broad differential between the 'A' and 'B' groups. There is no support, however, for the range of personality characteristics suggested by Morrisby for each individual profile.

Close observation of the subjects during the administration of the Differential Test Battery suggested that signs of personality were evidenced during the working of the so-called ability tests. It was decided to examine the correlation coefficients between these and the fourteen personality characteristics rated by the Training Officers. For good measure the two skill tests were also included. Correlation coefficients shown x and—in Table 9 are positively or negatively significant at the 5 per cent. level. From Tables 8 and 9 it will be noted that there is apparently a much greater degree of relationship between the personality characteristics and the ability tests than between the former and the speed tests underlying the modal profiles.

TABLE 8

PERSONALITY CHARACTERISTICS MARKEDLY POSITIVE IN EACH OF THE MODAL PROFILES.

Characteristics	"A" Variations			"B" Variations		
	Normal	Scatter-brained	Insipid	Normal	Tear-away	Not* known
1. Free and easy	x	x	x		—	—
2. Operates in terms of rules		—	—	—	—	—
3. Imaginative	—	—	—	—	—	—
4. Co-operative	x	x	x	x	x	
5. Reliable	x	x	x	x		
6. Willing to accept criticism	x	x	x	x		—
7. Flexible in mind	—	—	—	—	—	—
8. Steady		x	—		—	—
9. Willing to accept responsibility		—	—	—	—	—
10. Amenable to authoritarian discipline	—	x			—	—
11. Rigid	—	—	—	—	—	—
12. Ambitious	x	—	—	—	—	—
13. Tactful	—	x	—	—	—	—
14. Active under most circumstances	x	x	x			

*Morrisby does not interpret this profile which is :

No. & Na.	xxx xxx
Pers.	xxxxxxxxxxxx xxxxxxxxxxxx
Wd. Fl.	xxxxxxxxxxxx xxxxxxxxxxxx
Id. Fl.	xxxxxx xxxxxx

TABLE 9

SIGNIFICANT CORRELATION COEFFICIENTS BETWEEN THE TRAINING OFFICERS' RATINGS ON THE FOURTEEN PERSONALITY CHARACTERISTICS AND THE MORRISBY ABILITY AND SKILL TESTS.

Characteristics	Ability Tests.			Skill Tests.				
	CST	GAT V	GAT N	GAT P	Shapes	Mech.	Mot. Sp.	Mot. Sk.
1. Free and easy in his general approach ..			x	x				
2. Operates in terms of rules	x			x				x
3. Imaginative	x	x	x	x		x		
4. Co-operative	x	x	x	x	x			x
5. Reliable	x	x	x	x				x
6. Willing to accept criticism.....				x		x		x
7. Flexibility of mind ..	x	x	x	x		x		x
8. Steady	x			x				x
9. Willing to accept responsibility	x	x		x	x	x		x
10. Amenable to authoritarian discipline	x	x		x	x	x		x
11. Rigid		—		—				
12. Ambitious	x	x	x	x	x	x		x
13. Tactful	x			x		x		x
14. Active under most circumstances	x		x	x	x			x

A similar exploration of the correlation coefficients between individual tests of ability and skill, on the one hand, and each of the eight dimensions of the Employees' Attitude-Dimensions Scale on the other, showed only three significant correlation coefficients out of a total of sixty-four, these being significant at the 5 per cent. level.

(iii) *The Factorial Analysis of the Differential Test Battery.*

Morrisby suggests that the Differential Test Battery is not susceptible to factor analysis, since it is based on a number of 'differentiated' tests which would not show any factorial overlap. Nevertheless, it seemed appropriate to attempt a factorial analysis of the battery after obtaining the inter-correlation matrix. The latter is shown in Table 10, where the individual product-moment coefficients are all based on the main sample of 124 subjects.

In Table 10 the diagonal cells have been filled with the test reliabilities reported earlier. Four factors were extracted from this matrix using the method of Principal Components. The rotated loadings are shown in Table 11, significance again being rated at the 5 per cent. level. The communalities (h^2) are also shown,

TABLE 10

INTERCORRELATION COEFFICIENTS WITHIN THE DIFFERENTIAL TEST BATTERY.

Tests	CST	GAT V	GAT N	GAT P	Shapes	Mech.	No. & Na.	Pers.	Wd. Fl.	Id. Fl.	Mot. Sp.	Mot. Sk.
C.S.T.62											
GAT-V42	.76										
GAT-N48	.47	.72									
GAT-P ..	.61	.50	.51	.87								
Shapes36	.30	.29	.48	.77							
Mech.33	.37	.08	.37	.37	.75						
No. and Na.	.18	.39	.51	.37	.14	.06	.82					
Pers.	-.07	-.04	-.19	-.15	-.13	.01	.27	.82				
Wd. Fl. ..	-.01	.09	.25	.08	-.06	-.16	.23	-.04	.90			
Id. Fl.06	.02	.15	.06	.01	.03	.03	.09	.17	.81		
Mot. Sp. ..	-.05	.13	.13	.05	-.12	-.03	.34	.01	.22	.23	.71	
Mot. Sk. ..	.11	.18	.22	.26	.10	.26	.24	-.14	.05	.06	.21	.44

TABLE 11

FINAL ROTATION OF FACTOR LOADINGS FOR THE DIFFERENTIAL TEST BATTERY.

Factors	I	II	III	IV	h^2
Tests					
C.S.T.74	-.30	—	—	.65
GAT-V69	—	—	.22	.53
GAT-N80	.25	—	—	.71
GAT-P80	—	—	.19	.71
Shapes56	-.42	—	—	.50
Mechanical32	-.50	.21	.56	.71
No. and Na.54	.48	-.28	.24	.66
Pers.	—	—	.74	—	.63
Wd. Fl.29	.62	—	-.26	.54
Id. Fl.19	.29	.65	—	.54
Mot. Sp.	—	.64	.26	.40	.64
Mot. Sk.	—	—	—	.78	.66

The factor pattern is not particularly clear, but individual elements were identified as follows :

- I. Non-Verbal Reasoning.
- II. Speed of Identification.
- III. Ideational Flexibility.
- IV. Mechanical Manipulation.

Thus, there appears to be a mixture of different types of factors.

Table 11 throws some further light on the many profiles outlined by Morrisby. Factor I suggests that the three GAT tests form a group of a coherent nature giving information about reasoning in various forms, though it throws no light on the four corresponding individual profiles suggested by Morrisby. Factor IV seems to support the 'Mechanic' Profile thereby inferring the 'Craft' Profile by implication. This factor also lends some support to the 'Skill' Profile, but suggests that this may be closely related to its 'Mechanic' counterpart. Factor II lends some support to the allied 'Speed' Profile but suggests that this is also related to other tests in the battery. Factor III is closely related to the modal profiles and seems closest to Type B—Tear-Away ! Perhaps its very existence is interesting, however, since factors in test batteries such as the present one would normally tend to be of the non-personality type.

The results of the factorial analysis would, therefore, seem to offer some support to the notions of 'profiles' as suggested by Morrisby. It may be, however, that the ability, vocational or skill profiles are, in fact, fewer in type and involve more tests in each case. In addition the one personality-type factor is insufficient to lend weight to the modal profiles and once again there seems insufficient evidence for the existence of these.

V.—DISCUSSIONS AND CONCLUSIONS.

Very little evidence is so far available on either the reliability or the validity of the Differential Test Battery. Table 1 above shows the range of retest reliabilities obtained on a small number of cases in the present investigation, and these can be regarded as indicative of a reasonable degree of consistency of measurement for the various individual tests. More information is needed about the reliability of this battery.

Considerable claims have been made for the validity of the Differential Test Battery for vocational selection purposes. These claims are largely in terms of ability, modal and skill profiles, these involving respectively Tests 1-6, 7-10, 11-12. Compared with the corresponding Training Officers' Ratings in the present investigation, the ability showed a higher validity than the other two groups, but this yielded no evidence concerning the ability profiles. For the latter it was necessary to examine critically the results of the factor analysis of the battery. As pointed out earlier, the four factors obtained showed some traces of certain of the claimed ability profiles, but on a broader and less precise level. Nevertheless, certain other basic profiles could be established by re-grouping tests in the battery along the lines of the factors obtained ; such re-grouping could lead to the construction of profile charts for use in vocational selection on a broad front. It will be noted in both the comparison with the Training Officers' Ratings and in the later factor analysis, the test GAT-N stands out. This test shows no significant correlation with the Training Officers' Ratings for 'Thinking in Numbers,' but is highly loaded with the 'non-verbal reasoning' factor. Thus, it would appear that this test might be more theoretical than practical in its orientation, and thus it fails to play its assigned role amongst the ability tests as a whole.

When the validity of the modal profiles was examined against the criterion afforded by the newly constructed Employees' Attitude-Dimensions Scale, the four individual tests underlying these profiles showed few significant correlations with the eight dimensions of the attitude scale. Thus, there was little evidence of relationships between the personality characteristics called into play in each case. Examination of the scores in the high and low attitude groups, (upper and lower third of sample by total score on the attitude scale), showed that only the Number and Name Checking and Ideational Fluency tests distinguished significantly between these groups. It might be assumed, then, that these two tests provide discriminating instruments for general attitude towards vocational affairs. The combination of these two tests is also underlined in Factor III of the factor analysis, but in neither case do these results appear to reflect any aspect of the modal profiles. When the subjects in the high and low attitude groups were each assigned to the most appropriate of the eight modal profiles, the general indication was that the groups of high and low attitude tended to differentiate between the "A" and "B" type profiles—or, in reverse, that the "A" and "B" type profiles are significant of good or poor general vocational attitudes. Nothing more could be inferred.

Little further evidence was gained by an elaborate comparison of the single and combined modal profile tests with the Training Officers' Ratings for the fourteen related personality characteristics. A consideration of the correlation coefficients between the single tests and the ratings for individual personality characteristics showed that only the Number and Name Checking test showed any appreciable number of significant coefficients. The two fluency tests showed none at all. There was certainly no pattern of significant correlation coefficients across the modal profile tests for any one personality characteristic, and certainly no corresponding matrix for any group. When the combined correlation coefficients were obtained for each profile between the single tests and the group of related personality characteristics as ascribed by Morrisby, again few significant coefficients were obtained, and there was no evidence to support these groups as underlying the various modal profiles. Indeed, when the subjects were each grouped under the most appropriate profile according to their test results, the underlying groups of personality characteristics indicated by the correlations with the Training Officers' Ratings were seen to vary widely from those ascribed by Morrisby, indicating no validity whatever for this notion of a range of personality characteristics related to each individual modal profile. As in the case of the attitude scale, however, a close inspection of the relations between the individual modal profile tests and the ratings of the personality characteristics does show a reasonably broad differential between the groups of characteristics underlying the "A" and "B" profiles, and suggests that the broad distinction between these two groups of profiles might be a useful pointer in vocational selection. Further work is needed to clarify this. The factor analysis throws no further light on the validity of the modal profiles since the one personality-type factor cannot be identified with any one of these. Nevertheless, the presence of this one personality-type factor does reinforce the importance of personality characteristics in the general field of testing, and does suggest the possibility of testing for these in the kind of differential battery now under investigation.

Similar examinations of the validity of the skill profiles leads again to little general support for their use as suggested by Morrisby. A comparison of the scores on these two tests with the Training Officers' Ratings for the corresponding abilities showed only two significant correlation coefficients. A further

comparison of the same scores with the Training Officers' Ratings for the fourteen personality characteristics showed that one test had no relationship with any characteristic and the other, a significant relationship with almost every characteristic in the list ! The factor analysis probably offers the most cogent evidence here. The fourth factor certainly demonstrates the existence of the 'Mechanic' Profile, and, by implication, that of the 'Craft' Profile. This is the sole emergence of a profile strictly in line with those outlined by Morrisby for use in connection with the Differential Test Battery.

As a by-product of the present investigation, it is of interest to note that both the ability tests in the Differential Test Battery and the eight dimensions of the Employees' Attitude-Dimensions Scale correlate more highly with the Training Officers' Ratings for the fourteen personality characteristics than do the modal profile or skill tests. It may be that either of these instruments might eventually yield better measures of 'intellectual personality' than the present series of speed tests.

From the evidence reported in this paper, it may be seen that the Morrisby Differential Test Battery has some contribution to make to the assessment of human potential or performance in the field of vocational selection. The individual tests are of reasonable reliability and of suitable standard of difficulty for subjects such as those used in this investigation. Their administration is clearly described within the manual and lies within the competence of most personnel officers or psychologists. Scores from groups of tests could be of value on the broad fronts of both ability and personality assessments. So far as this investigation shows, however, there is no basis for accepting the complex system of profiles based on these tests as having any major relevance to vocational selection. No underlying patterns were found to support either the existence of these profiles or their relevance to job suitability in terms of either abilities or personality characteristics. Nevertheless, there was some evidence to indicate that a much smaller number of profiles might be obtained by re-grouping sets of tests which would have a firmer relevance in the general field of vocational selection.

ACKNOWLEDGEMENTS.—The writers are greatly indebted to Mr. G. Bostock and his colleagues in The British Iron and Steel Federation Training Department, to Mr. L. B. Birch of the University of Sheffield, and to the Training Officers in the firms concerned, for providing facilities for this investigation and for much help in carrying it out. They also wish to record their thanks for financial assistance given by the Central Research Fund, University of London, and the United Arab Republic Education Bureau, London.

VI.—REFERENCES.

- ¹ KORNHAUSER, A. W. (1922) Business applications of a mental alertness test. *J. Pers. Res.*, 1, 103-121.
- ² YOAKUM, C. S. (1922). Basic experiments in vocational guidance. *J. Pers. Res.*, 1, 18-34.
- ³ SEASHORE, R. H. (1928). Stanford motor skill unit. *Psychol. Mon.*, 39, 51-65.
- ⁴ STENQUIST, J. L. (1923). *Measurement of Mechanical Ability*. New York: Col. Univ., Cont. to Education, 130.
- ⁵ COX, J. W. (1928). *Mechanical Aptitude*. London: Methuen and Co., Ltd.
- ⁶ FARMER, E. (1927). A group factor in sensory motor rests. *Brit. J. Psychol.*, 17, 327-334.
- ⁷ EARLE, F. M., and MACRAE, A. (1929). *Tests of Mechanical Ability*. London: Studies in voc. guidance, 3, Nat. Inst. of Ind. Psych.

- ⁸ EARLE, F. M., GAW, F., *et al.* (1930). *The Measurement of Manual Dexterity*. London : Studies in voc. guidance, 4, Nat. Inst. of Ind. Psych.
- ⁹ EL-KOUSSY, A. A. H. (1936). The visual perception of space. *Brit. J. Psychol. Mon. Supp.*, 20.
- ¹⁰ THURSTONE, (L. L. 1938). Primary mental abilities. *Psychometric Mon.*, 1.
- ¹¹ BINGHAM, W. V., and DAVIS, W. T. (1924). Intelligence test scores and business success. *J. App. Psychol.*, 8, 1-22.
- ¹² FRYER, D. (1931). *The Measurement of Interests*. New York : Holt and Co.
- ¹³ STRONG, E. K. (1943). *Vocational Interests of Men and Women*. Stanford, U.S.A. : Stanford Univ. Press.
- ¹⁴ MAIER, N. R. F. (1955). *Psychology in Industry*. U.S.A. : Houghton Mifflin Co.
- ¹⁵ HULL, G. L. (1928). *Aptitude Testing*. New York, U.S.A. : World Book Co.
- ¹⁶ EL-SHARKAWY, M. K. I. (1963). *A Study of the Reliability and Validity of a Differential Test Battery in Relation to Industrial Training*. MA.. Thesis, University of London, 101-109.
- ¹⁷ MORRISBY, J. R. (1955). *Differential Test Battery*. London : National Foundation for Educational Research.
- ¹⁸ MORRISBY, J. R., *loc. cit.*
- ¹⁹ EL-SHARKAWY, M. K. I., *loc. cit.*, 196-218.
- ²⁰ FERGUSON, G. A. (1959). *Statistical Analysis in Psychology and Education*. New York : McGraw-Hill Book Co., 248-249.

(Manuscript received 10th October, 1964.)

RESEARCH NOTES

SPECIAL EDUCATIONAL TREATMENT IN THE ORDINARY SCHOOLS

By H. G. ARMSTRONG

(*Educational Psychologist, West Riding*)

In a recent research, Williams demonstrated that a disproportionate number of children attending special schools for the E.S.N., in South Wales, had Summer birthdays. He concluded that this was probably the effect of educational organisation, although other factors may have been operating. Orme has shown that in a group of severely subnormal children low intelligence is associated with Winter and Spring births and has suggested that changes in climatic temperature during pregnancy may be one factor in the relationship between season of birth and later intellectual performance. On the other hand, Davies, more recently, found that season of birth had no effect on a population of normal adults.

The tables below refer to children who have failed to make progress in their own schools and have been examined by the School Medical Officer to consider their suitability for admission to a special school but have been recommended for special educational treatment in the ordinary school. Table 1 gives the total number put forward by all schools in the West Riding during the school year 1963-64. This number is likely to be an under-estimate because some areas do not have access to a special school; other factors affecting the total are variations in assessment and lack of uniformity in making returns to the Central Office.

TABLE 1
QUARTERLY DISTRIBUTION.

First Quarter	June, July, August	81
Second Quarter	March, April, May	83
Third Quarter	December, January, February	50
Fourth Quarter	September, October, November	44

It can be reasonably assumed, in the absence of the quarterly proportions from which these figures have been drawn, that Spring and Summer-born children produce a greater number of learning failures than Autumn-born children and that more children from the younger half of the age group are to be found in special schools for the educationally subnormal than one would expect.

TABLE 2
SEX DIFFERENCES.

	Boys	Girls	Total
First Quarter	59	22	81
Second Quarter	54	29	83
Third Quarter	33	17	50
Fourth Quarter	37	7	44
TOTAL	183	75	258

Table 2, showing the high proportion of boys failing to make progress in the ordinary school as compared with girls, reflects the different rates of maturation between the sexes. By inference the actual number of younger children of both sexes to be found in special schools is likely to exceed expectation.

TABLE 3
DISTRIBUTION OF INTELLIGENCE.

	Mean I.Q.	Standard Deviation
First Quarter	83	10
Second Quarter	82	10
Third Quarter	81	12
Fourth Quarter	82	12

Table 3 shows that children considered as educationally subnormal because of failure to make progress in the ordinary school overlap the special school field at one end of the range and the normal ability group at the other. Because of 'test-sophistication' it is probable that the true mean lies near to the upper limit for admission to special schools.

TABLE 4
SCHOOLS OF ORIGIN.

Infants	J.B.	J.G.	J.M.	J.M. and I.	Sec. Mixed	Sec. B.	Sec. G.	Others	Total
13	8	5	80	72	62	6	6	6	258

That more than two-thirds of the children should be referred at the primary stage with a preponderance from junior mixed and junior mixed and infants is to be expected.

It is difficult to point to any single source of variation which can account for the unbalance demonstrated here. The three main sources of variation which make the situation more complex are :

- (i) Length of infant school life.
- (ii) School organization.
- (iii) Rate of development.

Observation suggests that younger children whose intellectual development is slow spend a shorter time in the infants' school than the older children and on transfer to a junior school are likely to be placed in a streamed group where motivation and stimulation are at a minimum.

There is a widespread feeling amongst teachers that all children should be admitted at the beginning of the school year, the inference being that, if they are exposed to the same amount of teaching, differences in maturation are likely to be overcome. That such an assumption is well-founded is difficult to accept. Table 2 shows clear differences in rates of development between boys and girls and, common to both, a disparity in the failure rate between the younger and older children. It is probable that the major factor is one of maturation and that, under present

circumstances, to admit all children at the beginning of the school year could exacerbate rather than improve the position, for the infant school would require to provide an environment similar to that of the nursery school and, at the same time, adequately meet the needs of those children more advanced in development.

Some modification of the present organisation is called for, and the one that is likely to have the greatest effect on the children discussed here is an extension of the nursery school environment, providing a rich variety of sensory experiences in order to extend their limited range of experience. In the junior school an extension of the most enlightened infant school practices, with careful nurturing of slow learning children in small remedial groups would also be valuable. It is not unreasonable to suppose that with an improved school environment there would be fewer casualties amongst the younger children and a more balanced entry to schools for the educationally subnormal. Further research, however, is needed to ascertain the nature of learning failure, differential rates of development and the incidence of the emotional disorders amongst children whose adjustment to the ordinary school environment is unsatisfactory.

REFERENCES.

- DAVIES, A. D. M. (1964). Season of birth, intelligence and personality measures. *Brit. J. Psychol.*, 55, 4, 475-476.
 ORME, J. E. (1963). Intelligence, season of birth and climatic temperature. *Brit. J. Psychol.*, 54, 273-6.
 WILLIAMS, P. (1964). Date of birth, backwardness and educational organisation. *Brit. J. Educ. Psychol.*, 34, 247-255.

(Manuscript received 11th March, 1965.)

A NEW PERSONALITY TEST FOR BOYS

By H. B. GIBSON

(The Institute of Criminology, University of Cambridge)

SUMMARY. This test was devised to overcome two obstacles to the testing of boys at an age when many are only semi-literate; (i) the limited reading and verbal capacity of such children, and (ii) the tendency of subjects to be reticent about personal problems. To overcome (i), the items of a simply worded personality inventory were printed separately on cards and illustrated with pictures; to mitigate (ii), the test was heavily buffered with items providing a check on veracity of response, and the material was presented as though it were a test of reading and comprehension.

The test was administered individually. Analysis of the results obtained from 285 boys aged 8 to 10 years, revealed two principal components, one explicable in terms of neuroticism and the other in terms of extraversion. Two independent scales were formed representing these personality dimensions.

Some preliminary validating data is available from teachers' ratings of the boys' behaviour, scores on the body sway test, and parental interviews by psychiatric social workers.

I.—INTRODUCTION.

Children's personality inventories are used both clinically, to assess the nature and degree of personal maladjustment, and experimentally, in the establishment of group differences with regard to variables such as delinquency (Cass, 1952; Hathaway and Monachesi, 1953; Simpson, *et al.*, 1960), deceit (Hartshorne and May, 1928; Gibson, 1964a), neuroticism, anxiety and extraversion (Himmelweit and Petrie, 1951; Casteneda, *et al.*, 1956; Sarason, *et al.*, 1960; Gallard and Goodfellow, 1962), social maturity (Findlay, 1955), and authoritarianism (Lyle and Levitt, 1955).

The authors of such studies are generally content with the 'construct validity' of the tests they use, that is, the test users accept the nomenclature employed by the test creators without questioning whether a high score on a variable like 'neuroticism' implies that low scorers are always 'more stable' in a meaningful sense of the term. Thus, Furneaux and Gibson (1961), in publishing the standardization of the J.M.P.I. claimed no more than that, "These (scales) facilitate the coding, as scores, of two kinds of statements made by children. The assertion that they are concerned with neuroticism and extraversion rests at the present time almost entirely on the face validity of the items in terms of which they are defined." But a number of users of this Inventory have tended to reify the terms 'extraversion' and 'neuroticism' as though they were taking measures as definite as height and weight. The later addition of a lie scale to the J.M.P.I. (Gibson, 1964a) has indicated that with this Inventory, as with many others, while a high neuroticism score may mean a high degree of neuroticism, a low neuroticism score sometimes merely reflects the fact that the subject is unwilling to be frank about himself.

The incorporation of lie scales into personality tests improves their validity, but only at the expense of discarding the scores of a proportion of the population judged to be 'liars' (of whom, some will be telling the truth). But even when the records of the 'liars' have been discarded, a proportion of the remainder may be quite invalid because of the limited comprehension of some of the subjects, for random responses may lead to quite average scores on all scales. While the problem of test-faking confronts all psychologists, the child psychologist is especially concerned with the problem of test-comprehension as well. Although children of all ages will cheerfully tick 'yes' or 'no,' 'same,' or 'different,' when they have been shown how to respond to a self-rating instrument, this is no guarantee that their responses are related to the *meaning* of the items even when these are read aloud to them. Yet, psychologists still give self-rating instruments to children as young as six years (e.g., Rie, 1963). Those who have first-hand experience of teaching and testing children in infant and junior schools will fully appreciate the artefacts that can arise from children's readiness to respond at random rather than to admit ignorance.

The present study concerns the construction of a new children's personality test under the pressure of necessity, and investigation of its validity by three independent criteria of personal adjustment. This test gives special attention to the two problems of test-faking and incomprehension, because the population studied was mainly composed of boys aged 8 to 9 years with a poor educational background. Also, there was a special interest in the types of boy who might be likely to resist revealing themselves too frankly to a visiting psychologist.

II.—TEST MATERIAL.

The test material consisted of forty-two cards, 3 × 5 inches in size, and two posting boxes. On each card there was a printed statement and a very simple picture illustrating the meaning of the statement. For example, one statement read: "He is not afraid of big dogs," and the drawing depicted a smiling boy patting a large dog. The statements on the cards were worded very simply, and described the appearance, habits, likes and dislikes of a boy about 8 to 10 years old. When compiling the list of items, social workers, teachers and others were consulted so that the phrases used would be entirely meaningful to boys from working-class homes.

There were three categories of cards: (a) fourteen cards bearing statements of a cheerful nature (as above); (b) fourteen cards describing unpleasant situations or feelings (e.g., "Other boys bully him"); (c) the remainder having statements with a little emotional content. These latter cards served not only as buffer items, but as a means of training the subject to respond promptly and truthfully. Their content was such that the experimenter could generally check a 'wrong' response immediately, as explained below.

One of the posting boxes bore the words "Same as me," and the other "Not the same."

III.—ADMINISTRATION.

The test was presented to each boy individually, as though it were a test of reading and comprehension. The experimenter explained that the task was to read the sentence on each card out aloud, and then to post it in the box appropriate to its meaning. Thus, when the subject was presented with the card bearing the statement, "He has red hair," the experimenter checked that the boy put it in the proper box, according to the colour of his hair. As most of the buffer items contained objectively verifiable statements, the subject could be checked for any tendency to respond other than strictly according to the truth about himself on these items. To encourage a quick and spontaneous response to each card, the experimenter said that he wanted to see how long it would take the subject to get through the pile of cards, and left a stopwatch, running in full view.

Where the boy was a very poor reader, it was sometimes necessary to help with the reading, and such assistance was readily given. As the meaning of the statements was conveyed by the pictures on the cards, poor readers could often succeed by guessing correctly. The cards were presented in a fixed order, so that one from each category came in turn.

Subjects.

The test was administered to a total of 389 boys aged 8 to 10 years. All subjects were seen in individual sessions during which other psychological tests were given.

IV.—RESULTS.

The records of 285 boys were submitted to a principal component analysis.* These were the boys who happened to be tested first, with a minority of cases excluded where the quotient on reading or verbal comprehension was less than 76. The first three components extracted accounted for 21.02 per cent. of the total variance.

In order to arrive at a simple structure, the positions of all the items (except buffer items) were plotted with respect to their loadings on the two orthogonal axes of the second and third components (C_2 and C_3). On this graphical plot, a new axis (C'_2) was drawn at 70° to C_3 , and the loadings of the items on this new axis were calculated. On a fresh graph, C'_2 was drawn orthogonal to the axis of the first component (C_1), and the positions of all the items were plotted with respect to their loadings on these two axes.

There were two recognisable clusters of items. The cluster with high loadings on C_1 , contained items with unpleasurable content such as "His Mum picks on him too much," and "He often feels poorly when it is time to go to school." Endorsement of such items implies complaining about unpleasurable experiences and mistreatment by others, so it seems reasonable to interpret it in terms of neuroticism. The other cluster contained the cheerful-sounding items, e.g., "He never worries." These two clusters were not diametrically opposed to one another, but were almost orthogonal. It was expected that two such items as "He never worries," and "He has lots of worries," would be diametrically opposed, but this was not the case. It appears that the endorsement of a cheerful-sounding statement is not the opposite of endorsing a gloomy-sounding statement, and that psychological attitudes which are different in kind are responsible for these two clusters. If the cluster high on C_1 is to be labelled 'neuroticism' then that high on C'_2 cannot reasonably be labelled 'stability.' Although in designing the test, the experimenter's interest was in maladjustment, the personality factor of rathymia or extraversion has emerged from this analysis, for endorsement of the cheerful-sounding items can probably best be explained by this personality attribute.

In order to form a neuroticism scale, the items with loadings over .4 in the C_1 cluster were given a weighting of two points and the others of one point. The extraversion scale was formed by scoring one point for each of the thirteen items in the other cluster. That double weightings were used in one scale but not in the other is justified by the fact that the former was the major component accounting for the highest percentage of the total variance.

*This analysis was carried out by the University of London Computer Unit.

Normative data.

The test was administered to 389 boys in all, and scored for the two scales of neuroticism and extraversion. Excluding the records of boys whose quotients on reading or verbal comprehension were less than 76 (on grounds of doubtful validity) the distributions of scores are shown in Table 1. Both distributions are highly skewed, but in opposite directions. Because of this asymmetry it is not meaningful to express norms in terms of means and standard deviations. Instead, the percentile levels of the scores are shown. There is no evidence of any significant correlation between the two scales ($r = -.07$).

V.—VALIDATION.

Proper validation of this test must await studies on populations other than the population with which the test was constructed. There are some preliminary data available, however, which comes from three separate sources: (a) teachers' ratings on a scale described elsewhere (Gibson, 1964b); (b) data collected about the boys by psychiatric social workers who interviewed their families and checked with child guidance clinics, care committees, etc.; (c) the body sway test. The general nature of this information indicates that boys who are behavioural problems, being rated as 'naughty' by their teachers, and as suffering from 'conduct disorder,' in terms of P.S.W. ratings, tend to have significantly higher scores on both neuroticism and extraversion. The neuroticism scale alone does not discriminate so significantly. Such boys also tend to exhibit high static ataxia and sway on the body sway test as administered in this study after the manner described by Eysenck (1947).

VI.—REFERENCES.

- CALLARD, M. P., and GOODFELLOW, C. L. (1962). Neuroticism and extraversion in school-boys as measured by the J.M.P.I. *Brit. J. Educ. Psychol.*, 32, 241-250.
- CASS, LORETTA, K. (1952). Parent child relationships and delinquency. *J. Ab. Soc., Psychol.*, 47, 101-104.
- CASTENDA, A., et al. (1956). The children's form of the manifest anxiety scale. *Child Developm.*, 27, 317-326.
- EYSENCK, H. J. (1947). *The Dimensions of Personality*. London: Routledge and Kegan Paul.
- FINDLAY, C. B. (1955). The social opinions inventory as a measure of social maturity. *Child Development*, 26, 81-90.
- FURNEAUX, W. D., and GIBSON, H. B. (1961). A children's personality inventory designed to measure neuroticism and extraversion. *Brit. J. Educ. Psychol.*, 31, 204-207.
- GIBSON, H. B. (1964a). A lie scale for the Junior Maudsley Personality Inventory. *Brit. J. Educ. Psychol.*, 34, 120-124.
- GIBSON, H. B. (1964b). The spiral maze: a psychomotor test with implications for the study of delinquency. *Brit. J. Psychol.*, 54, 219-225.
- HARTSHORNE, H., and MAY, M. A. (1928). *Studies in Deceit*. New York: Macmillan.
- HATHAWAY, S. R., and MONACHESI, E. D. (1953). *Analysing and Predicting Juvenile Delinquency with the M.M.P.I.* Minneapolis: Minnesota U.P.
- HIMMELWEIT, H. T., and PETRIE, A. (1951). The measurement of personality in children. *Brit. J. Educ. Psychol.*, 21, 9-29.
- LYLE, W. H., and LEVITT, E. E. (1955). Punitiveness authoritarianism and parental discipline of grade school children. *J. Abnorm. Soc. Psychol.*, 51, 42-46.
- RIE, H. E. (1963). An exploratory study of the C.M.A.S. lie scale. *Child Development*, 34, 1003-1017.
- SARASON, S. B., et al. (1960). *Anxiety in Elementary School Children*. New York: John Wiley and Sons.
- SIMPSON, J. E., et al. (1960). Delinquency potential of pre-adolescents in high delinquency areas. *Brit. J. Delinquency*, 10, 211-215.

(Manuscript received 15th June, 1964.)

TABLE 1

FREQUENCY DISTRIBUTIONS OF SCORES OF 305 BOYS ON NEUROTICISM AND EXTRAVERSION.
(Boys with quotients < 76 on reading and verbal comprehension have been excluded).

Score	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Neuroticism: ..																		
Cum. frequency	22	42	75	111	148	180	212	233	251	271	277	289	295	299	300	304	305	305
Percentiles	7.2	13.8	24.6	36.4	48.5	59.0	69.5	76.4	82.3	88.9	90.8	94.8	96.7	98.0	98.4	99.7	100	100
Extraversion:																		
Cum. frequency	0	0	0	1	5	11	21	54	82	129	207	262	294	305				
Percentiles	0	0	0	0.3	1.6	3.6	6.9	17.7	26.9	42.3	67.9	85.9	96.4	100				

THE GUESSING CORRECTION IN VOCABULARY TESTS

By M. HAMMERTON

(Medical Research Council, Applied Psychology Research Unit, Cambridge)

The 'guessing correction' usually applied to subjects' scores in multi-choice vocabulary tests is examined, and doubt is cast upon its validity. Two experiments are described; the first was inconclusive; but the second strongly indicates that the guessing correction is inadequate. It is concluded that it can pay to guess.

INTRODUCTION.

A very common form of vocabulary test is that wherein the subject is presented with a word and a group of n other words, from which he has to select the one nearest to the word first presented. It is clear that a purely random response would, on the average, give the correct answer $\frac{1}{n}$ of the time, and the wrong answer $\frac{n-1}{n}$ of the time. It is usually assumed that errors arise from guesses, and that x errors will be made in $\frac{xn}{n-1}$ guesses. Consequently, a 'correction' is made to a subject's raw score. If he has made m correct and x incorrect responses, his final corrected score is $(m - \frac{x}{n-1})$.

However, the unstated assumption here is that the subject has only two response states: certainty and blind guessing. It is mere common sense that this is hardly the case. In two analogous situations, moreover, it has been shown that such an assumption leads to error. Broadbent and Gregory (1963) showed how the findings of signal detection experiments could only be understood if this assumption were abandoned; and Brown (1965 a, b) has similarly demonstrated its inadequacy in the field of immediate memory.

It therefore seemed reasonable to enquire whether the 'guessing correction' does, in fact, correct. Common sense would suggest that, if much 'guessing' were based upon a response state, intermediate between certainty and randomness, the standard correction would be inadequate, and that it would pay to guess. Two experiments were devised to see whether this was so.

EXPERIMENT 1

Method.

It was decided, as a first step, to administer a vocabulary test to a group of subjects who would be able to give, with each response, their own assessment of its reliability. The instruction to the subjects was, therefore, modified in the following way. Subjects were, as usual, requested to underline the word they thought appropriate; but were asked to use a triple underlining if they were entirely confident that they were right, double underlining if they were in doubt, and a single line if they were very doubtful.

The subjects used—twenty in number—were the members of an evening institute class to which the writer was lecturing. The test used a modified version of the vocabulary part of Shipley's self-administering test (Shipley, 1940). As it seemed that this might be too easy in its original form, the last ten odd numbered words were deleted, and ten more obscure words were added to the end.

Results and Discussion.

As was reasonably predictable, the percentage of correct choices declined with confidence (see Table 1). However, though t -tests show that the mean percentage correct of the 'entirely confident' category was significantly greater than that in the 'some doubt' category ($p < .005$), which in its turn was greater than that in the 'very doubtful' category ($p < .05$), the mean percentage correct in the 'very doubtful' category was just not significantly greater than 25 at the 5 per cent. level.

Nevertheless, the tendency was in the expected direction; and there were many reasons for accepting these rough figures with caution. For example, not all the subjects used all the categories; and the number of omissions was very various. A further experiment was, therefore, devised and carried out.

TABLE 1
PERCENTAGE OF CORRECT CHOICES IN EACH CONFIDENCE CATEGORY.

Category	Very confident	Some doubt	Very doubtful
Percentage correct	95.2	49.7	35.4

EXPERIMENT II.

Method.

As a preliminary, the split-half correlation of the scores obtained by subjects in experiment I—i.e., the correlation between scores obtained with the even and with the odd numbered questions—was obtained. Pearson's r value for this was $+0.87$, which is very highly significant ($p < .001$). It was, therefore, decided to use the odd and even numbered questions as two separate tests, hereinafter called tests A and B, which could be regarded as equally difficult and giving strictly comparable scores.

These tests were used with two differing sets of instructions. Instruction 1 was to leave out those questions where the subject did not feel confident of his knowledge; instruction 2 was to answer all the questions, and to guess in the absence of sure knowledge.

The subjects used were twenty-four adult volunteers. These were divided into two groups of twelve each, balanced as far as possible for scores on a non-verbal intelligence test. One group was given instruction 1 first, followed by instruction 2; the other was given the instructions in reverse order. Both groups were further sub-divided into groups of six, one of which had test A, followed by test B, the other having test B, followed by test A. The design was thus balanced for order of presentation of tests and of conditions.

Results.

Each subject had obtained two scores (both subjected to the standard 'guessing correction') one with and one without encouraged guessing. With guessing, twelve subjects showed improvement, eight exhibited no change, and four were worse. The mean score was 12.4 with guessing and 11.8 without, the respective standard deviations being 4.45 and 4.55. As the tests had been shown to be comparable, each subject's scores constituted a matched pair, and the Wilcoxon signed rank test was used to investigate the significance of this difference. It was, in fact, found to be significant at the 2 per cent. level (2-tail).

DISCUSSION.

These findings seem to confirm the suggested doubts of the validity of the standard guessing correction. Whilst it remains in use, the recommended strategy for examinees is to guess if they do not know. More seriously, the logical next step would be to seek a method of finding the guessing characteristics of the testee in a manner analogous to those now common in signal detection experiments. This, however, clearly presents great difficulties; and at the present time no such method is in sight.

REFERENCES.

- BROADBENT, D. E., and GREGORY, M. (1958). Division of attention and the decision theory of signal detection. *Proc. Roy. Soc., B.*, 158, 222-231.
- BROWN, J. (1965). Multiple response evaluation of discrimination. *Brit. J. Statist. Psychol.* In press.
- BROWN, J. (1965). A comparison of recognition and recall by a multiple response method. *J. Verb. Learn. and Verb. Behaviour.* In press.
- SHIPLEY, W. C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. *J. of Psychol.*, 9, 371-377.

(Manuscript received 25th January, 1965.)

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES

THE RELATIONSHIP BETWEEN READING READINESS AND READING PROGRESS

By D. V. THACKRAY

(Summary of thesis presented for the degree of M.A. in Education, in the University of London, 1964.)

METHOD.

Research in the field of reading readiness suggests that the important factors involved in reading readiness are, specific reading readiness skills such as visual and auditory discrimination, mental ability, home environment and emotional and personal attitudes.

A representative sample of 182 children from eleven schools were tested in these factors of reading readiness when commencing their second term in school (average age 5 years, 4 months).

The following measures were used for :

(1) *Reading readiness skills.*

The Harrison-Stroud Reading Readiness Profiles (anglicised by the writer) consisting of using symbols, making visual discriminations, using the context, making auditory discriminations, and using context and auditory clues.

(2) *Mental ability.*

(i) The Kelvin Measurement of Ability Test for Infants.

(ii) Teachers' ratings of general ability.

(3) *Home environment.*

(i) A multiple-choice Picture Vocabulary Test constructed by the writer.

(ii) Teachers' ratings of language and speech.

(iii) Notes made by the writer on the socio-economic background of the children's homes.

(4) *Emotional and personal attitudes.*

Class teachers' ratings of self-confidence, co-operation with adults, co-operation with other children, persistence, stability, and prevailing attitude.

Two terms later the same children (average age 6 years) were tested for reading achievement by the use of the Southgate Group Reading Test and the 'primer criterion.' The children were tested again with these measures at the beginning of their fifth term in school (average age 6 years 4 months) to examine reading progress.

ANALYSIS OF DATA.

- (1) A study was made of the competency of the children tested on all the aspects of reading readiness.
- (2) The raw scores obtained from all the measures used were standardised, and all the earlier results from the readiness measures were correlated with the two later reading achievement results. The predictive value of the reading readiness tests and the relative importance of the various aspects of reading readiness were assessed by the degree of correlation between these two sets of scores.
- (3) A comparison was made between the performances of the boys and girls on both the reading readiness measures and the reading achievement tests.
- (4) The minimum mental age requirements for effective reading progress in the sample, was investigated.

RESULTS.

The principal results are summarised as follows :

- (1) The frequency distribution of scores on the five tests of reading readiness skills and also the measures of vocabulary, language and speech showed a wide range of individual performances. The mental age range was from 4.0 to 8.0 years. Teachers' ratings of emotional and personal attitudes tended to cluster round the average.
- (2) Table 1 shows the important correlation coefficients obtained in this investigation. In this sample, the Harrison-Stroud Reading Readiness Profiles, as a whole, proved a valid measure of readiness for reading (.59), and of these tests, auditory (.53) and visual discrimination (.50) correlated the most highly with reading achievement; general ability (.47) and home environment (.42) were found to be important, but emotional and personal attitudes relatively unimportant (.10—.36).

TABLE 1
CORRELATIONS BETWEEN THE MEASURES OF READING READINESS AND THE MEASURES OF READING ACHIEVEMENT.

Measures of Reading Readiness	1st Reading Test (Southgate A)	2nd Reading Test (Southgate B)
	Correlation	Correlation
Harrison-Stroud Reading Readiness Profiles :		
Using symbols417	.400
Visual discrimination500	.489
Using the context341	.411
Auditory discrimination529	.531
Using context and auditory clue403	.433
Total score592	.596
General Ability :		
Kelvin Measure of Ability444	.409
Teachers' ratings of general ability506	.524
Home Environment :		
Vocabulary Profile431	.389
Teachers' ratings of language492	.468
Teachers' ratings of speech386	.406
Emotional and Personal Attitudes :		
Self-confidence337	.273
Co-operation with adults341	.287
Co-operation with children103*	.134*
Persistence361	.353
Stability219	.165
Prevailing attitude356	.332

* Not significant. All others significant at .1 per cent. level.

- (3) Table 2 shows the separate performances of boys and girls on the reading readiness measures and the reading achievement tests. The girls showed a significant superiority over the boys in two of the tests of reading readiness skills, viz., auditory discrimination (1 per cent. level) and using context and auditory clues (5 per cent. level)—the common element being auditory discrimination. Also, in the Kelvin Measurement of Ability Test, the Vocabulary Profile and the two reading achievement tests the girls showed a significant superiority over the boys.

TABLE 2

COMPARISON BY SEX OF THE RAW SCORES ON THE READING READINESS MEASURES AND THE READING ACHIEVEMENT TESTS.

Reading Readiness Measures	Boys (100)		Girls (80)		Diff. in means	Stand. error of diff.	t	Stat. Sign.
	Mean	S.D.	Mean	S.D.				
Harrison-Stroud Reading Readiness Tests :								
Using symbols	13.02	6.90	12.88	7.70	.14	1.1	.13	*
Visual discrimination . .	14.75	8.46	17.00	8.49	2.25	1.27	1.77	*
Using the context	11.78	3.64	12.68	3.18	.90	.5	1.8	*
Auditory discrimination	9.42	3.33	11.15	3.72	1.73	.53	3.26	1%
Using context and auditory clues	8.25	3.05	9.35	3.85	1.1	.52	2.11	5%
Total score	57.1	20.00	62.5	20.2	5.40	3.02	1.79	*
Kelvin Measurement of Ability	29.40	12.40	33.87	12.00	4.47	1.83	2.44	2.5%
Vocabulary Profile	30.35	10.05	33.55	10.25	3.20	1.52	2.10	5%
Reading Achievement Tests :								
1st Reading Achievement Test	11.63	6.24	14.97	7.5	3.34	1.04	3.21	1%
2nd Reading Achievement Test	12.98	7.38	17.3	8.55	4.32	1.21	3.57	.1%

* Not significant.

- (4) In the sample, a mental age of 5 years 6 months was found adequate for beginning to read and for making satisfactory progress in the early stages of reading.

(Manuscript received 27th November, 1964.)

THE GROWTH OF LOGICAL THINKING IN NORMAL AND SUB-NORMAL CHILDREN

BY STEPHEN JACKSON

(Abstract of thesis submitted for the degree of M.Ed. in the University of Manchester, April, 1963)

AIM.

By administering several of Inhelder's experiments to English children of (a) normal intelligence, (b) subnormal intelligence, it was hoped to discover how far the growth of logical thinking, as described in Piaget and Inhelder's *The Growth of Logical Thinking from Childhood to Adolescence*, might be revealed in these two groups.

EXPERIMENTS.

The experiments were those described in the above work as Conservation of Motion, Equilibrium in the Balance, Falling Bodies on an Inclined Plane, Communicating Vessels, The Law of Floating Bodies, and Oscillation of a Pendulum.

SUBJECTS.

The subjects were (a) forty-eight children with I.Qs. between 90 and 110, comprising four boys and four girls in each of the age-groups 5, 7, 9, 11, 13 and 15, and (b) forty children with I.Qs. between 60 and 80 comprising four boys and four girls, in each of the age-groups 7, 9, 11, 13 and 15.

All the (a) group were drawn from one primary school and one secondary school. The (b) group were drawn from E.S.N. day schools, thirty-six from one school and four from two other schools. The majority of the children were from families of low socio-economic status in the Manchester area.

ADMINISTRATION.

Each child was interviewed individually and completed the six experiments in one sitting, a maximum of half-an-hour being allowed for each experiment. The type of interview was a compromise between the 'clinical' and the standardised, a type of set questions being supplemented at the discretion of the questioner. On the completion of the battery of experiments, each child was given Raven's Progressive Matrices.

The criteria used in assessing responses were taken from the book quoted above. In pooling the results of the six experiments, an overall sum of 'logical thinking' was calculated for each child by awarding 1 mark for each sub-stage of each test. (Thus, a child at sub-stage 1A on any test was given a sum of 1 for that test, a child of sub-stage 1B being given a sum of 2, etc.) The results were plotted on Guttman-type scalograms and rank correlations were calculated between logical thinking scores and (a) C.A. and (b) Raven's Matrices scores.

RESULTS.

TABLE 1

(Normal subjects.)

NUMBER OF SUBJECTS IN EACH AGE-GROUP REACHING EACH SUB-STAGE.

Ages	STAGES					
	1A	1B	2A	2B	3A	3B
5	14	30	4	—	—	—
7	2	23	23	—	—	—
9	—	4	21	22	1	—
11	—	—	9	36	3	—
13	—	—	5	22	19	2
15	—	—	1	23	19	5
Totals.....	16	57	6	103	42	7

TABLE 2

(E.S.N. Subjects.)

NUMBER OF SUBJECTS IN EACH AGE-GROUP REACHING EACH SUB-STAGE.

Ages	STAGES					
	1A	1B	2A	2B	3A	3B
7	12	36	—	—	—	—
9	6	22	18	2	—	—
11	—	26	22	—	—	—
13	—	17	27	4	—	—
15	2	18	24	—	—	—
Totals.....	20	119	91	10	—	—

TABLE 3

(Normal Subjects.)

NUMBER OF SUBJECTS REACHING EACH SUB-STAGE IN EACH EXPERIMENT.

Experiments	STAGES					
	1A	1B	2A	2B	3A	3B
Cons. of Motion	—	11	5	22	10	—
Oscill. of Pendulum	10	5	5	12	11	5
Communicating Vessels ..	—	9	22	12	4	1
Falling Bodies	—	12	11	17	8	—
Equilibrium in Balance ..	1	13	8	19	6	1
Law of Floating	5	7	12	21	3	—
Totals	16	57	63	103	42	7

TABLE 4
(E.S.M. Subjects.)

NUMBER OF SUBJECTS REACHING EACH SUB-STAGE IN EACH EXPERIMENT.

Experiments	STAGES					
	1A	1B	2A	2B	3A	3B
Conserv. of Motion	—	23	16	1	—	—
Oscill. of Pendulum	6	29	4	1	—	—
Communicating Vessels	—	23	17	—	—	—
Falling Bodies	—	15	21	4	—	—
Equil. in the Bal.	8	8	22	2	—	—
Law of Floating	6	21	11	2	—	—
Totals	20	119	91	10	—	—

It is necessary to point out that for the experiments on Conservation of Motion, Communicating Vessels, and Falling Bodies Piaget does not divide stage 1 into two parts, and that in the above tables all stage 1 responses have been recorded as at stage 1B.

It can be seen that in three experiments no subject in the normal sample reached stage 3B, which was reached by Piaget's subjects between the ages of 12 and 16. Table 4 shows that 95 per cent. of E.S.N. responses were no higher than the first stage of concrete operations, stage 2A : and that no subject reached stage 3.

In addition to numerical data the thesis contains extracts from protocols to show the quality, rather than the logical level, of the responses at every sub-stage and for every age group. These extracts, in revealing intermediate stages and the wide variety of mental models used, amply justify the use of the clinical method.

Other researchers have suggested that a subject's performance may be affected by the content of a task as distinct from the level of reasoning required. The present study supports this finding and Tables 5 and 6 show the range of logical-thinking levels reached by individual subjects in their six tests.

TABLE 5
RANGE OF INDIVIDUAL SUBJECTS' SIX SCORES (NORMAL SAMPLE).

	No. of substages included in range			
	1	2	3	4
Percentage of total Subjects	10	60.4	22.9	4.2

TABLE 6
RANGE OF INDIVIDUAL SUBJECTS' SIX SCORES (E.S.N. SAMPLE).

	No. of substages included in range		
	1	2	3
Percentage of total subjects	10	62.5	27.5

Using the two samples as one group, rank correlations were calculated as follows: logical-thinking and C.A., +.608; logical-thinking and Raven's Matrices scores, +.859. There was no significant difference between the scores of boys and those of girls ($.9 > p > .50$).

Among the 5-year-olds and E.S.N. 7-year-olds many responses were found at a lower level than those quoted by Piaget as illustrating Stage 1, suggesting that Piaget has reported only the higher levels of pre-operational thought. Many of the responses said by Piaget to be typical of certain stages were not found, or were found accompanying behaviour that belonged to a different stage. Small details of apparatus and questioning were found crucially to affect responses and, therefore, assessments.

Owing to the size of the samples and the subjective nature of the questioning and assessments, the findings of the present study cannot be taken as valid and reliable norms, but they may add to the growing body of data upon which generalisations may eventually be based.

MAIN CONCLUSIONS.

(1) The ability to think at a certain logical level in one situation does not necessarily imply the ability to think at that level in other situations.

(2) Apart from the failure of many older subjects to reach stage 3B, the responses of the normal sample generally confirm the ages at which Piaget suggests certain levels of logical thinking are reached. But this study also indicates that wide variation in levels of logical thinking may exist among children of similar mental age.

(3) E.S.N. children showed very little increase in scores beyond the age of 9 and there was a suspicion of deterioration between 13 and 15. Fifteen-year-olds achieved no higher level of logical thinking than did normal 8-year-olds.

(4) Piaget's definitions of stages of logical-thinking, in terms of children's comments and test-behaviour, appear to require more critical examination.

(Manuscript received 8th October, 1964.)

CRITICAL NOTICE

DOUGLAS, J. W. B. (1964). *The Home and the School: A Study of Ability and Attainment in the Primary Schools*. London: MacGibbon and Kee, pp. xxvi+190, 36s.

By CYRIL BURT

This book contains "the third main report on a follow-up investigation initiated by the Population Investigation Committee." The survey, we are told, "was originally undertaken to examine the availability and effectiveness of the ante-natal and maternity services." Dr. Douglas, who has directed the inquiry since it began in 1945, is by training a gynaecologist, and was formerly Reader in Public Health, at the University of Edinburgh. The investigation was subsequently prolonged in order to determine the conditions which influenced "the major types of illness suffered by the children up to the age of 5." Later still, it was decided to extend it for yet another period, with a view to studying, not (as might have been anticipated) the health and physical development of the children and their varying susceptibility to different types of ailment, but their "ability and attainment during their school years."

However, as Professor Glass explains in his Introduction, all the researches are "linked by a continuing interest in the influence of environment." The aims of the present inquiry were apparently twofold: first, to show how, even within the narrow span of primary school life, "measured intelligence responds to environmental factors," and secondly, to study how far environmental conditions affect the allocation of children to grammar and other secondary schools. This investigation, like the preceding, has been generously supported by grants from the Medical Research Council, the Nuffield Foundation, the London School of Economics, and several other bodies.

"In recent years," says Dr. Douglas, "the efficiency of our educational system has been the subject of intense concern"—a concern, he adds, clearly emphasized by the chief clauses in the Education Act of 1944. Hence, it has become especially important "to know how large the pool of ability would be if there were no leakage of talent." The "pool of ability" he defines as "the number of persons capable of benefiting from a university education." "Till now," he continues, "there has been no means of making an informed guess at its size." This is a somewhat surprising statement, since there have been many such inquiries carried out at various dates (see, for example, this *Journal*, XIII, 1943, pp. 83-98, where an estimate was given both for the size of the 'pool' and for the extent of the 'leakage'). The sample of children studied by the National Survey of Health and Development, so Dr. Douglas believes, supplies just the basis needed for answering the questions he has in view.

The sample consisted of over 5,000 boys and girls, all born in the first week of March, 1946, and "drawn from every type of family and every part of the country." To demonstrate his conclusions he has relied primarily on the scores obtained from two sets of tests administered at the ages of 8 and 11. Now, as Professor Hearnshaw has recently reminded us, "many of our most eminent psychologists have been, and still are, highly sceptical of mental testing: in their opinion the qualities which such tests profess to measure are not amenable to quantitative assessment, and, indeed, often do not exist, and the arguments based on such ephemeral results simply lend a spurious air of science to the writer's preconceived opinions." It, therefore, behoves the critical reader, before accepting the conclusions based on such procedures, to consider most carefully the snags and fallacies that beset inquiries of this type, and to observe how far the investigator has succeeded in avoiding them.

In selecting appropriate tests, Dr. Douglas had the advice of Dr. Pidgeon of the National Foundation of Educational Research, who briefly describes them in one of the appendices. With the 8-year-olds the following four tests were used: (a) for

'ability,' Stuart's 'picture test'; (b) for 'attainments,' a test (1) of mechanical reading, (2) of vocabulary (the same words for both), and (3) of 'reading comprehension.' There were no tests of arithmetic, spelling, writing, or drawing. For the 11-year-olds the following four tests were used: (a) for 'ability,' a combined verbal and non-verbal test; (b) for 'attainments' the same test of (1) mechanical reading and (2) vocabulary as before, and (3) a test of arithmetic. The administration of the tests was entrusted to the teachers. Of the original sample of 5,362 only 3,297, i.e., about three-fifths, actually took both sets. Of the rest some had died; others could not be traced; twins were for some reason excluded; and children who were "too backward to take the tests" were apparently omitted from the inquiry. As Dr. Pidgeon acknowledges, there is no direct evidence available to show that these particular tests do, in fact, furnish valid measures of the qualities they are intended to assess. Hence, if, as we are told, the results of the survey strongly suggest that environmental conditions affect the children's performance, that does not necessarily prove (as Professor Glass and Dr. Douglas would have us believe) that the children's 'intelligence' is affected by environmental conditions. The natural conclusion would rather be that the simple types of test employed cannot after all be regarded as supplying a trustworthy estimate of 'intelligence' in the sense in which the psychologist uses the term.

Before comparing the results, the marks from all four tests were averaged. This is a rather surprising procedure in a research which claims, in the words of the title, to be "a study of ability and attainment," and was planned to assess the extent to which 'ability' is wasted. The reason given is that, as two preliminary tables are said to show, "the circumstances which were associated with deterioration in the scores for achievement are equally associated with a deterioration in the mental ability scores": ('equal' certainly goes too far). Incidentally, the reader should be warned that, when the 'ability' of the 'underprivileged children' is said to 'deteriorate,' the word is used in a somewhat unusual sense. With increasing years the performances measured actually improve: all that is meant is that, with certain children, the amount of improvement is smaller than with others.

In his earlier investigations, Dr. Douglas followed the practice of most previous investigators and classified the children's home conditions according to the father's occupation: that, as he says, mainly determines income, and income determines food, housing, clothing, sanitation, and other factors influencing early development. For the present inquiry he decided to re-classify the homes on a new basis, determined partly by social 'background' and partly by the educational achievements of the parents. Four classes are distinguished: (1) *Upper Middle Class*: "One of the parents was brought up in a non-manual worker's family and went to a secondary school, and the other had at least one of these characteristics."; (2) *Lower Middle Class*: "Parents may be said to deviate in their upbringing from the middle class pattern"; (what is meant by 'deviation' is unexplained); (3) *Upper Manual Working Class*: "One or both parents come from a middle class family, or have been to secondary schools"; (4) *Lower Manual Working Class*: Both parents come from manual worker's families, and received an elementary education. Dr. Douglas remarks that the new classification displays a far more satisfactory agreement with the children's performances in his tests and in the examinations at 11 plus; indeed, he seems deliberately to have looked for criteria which would emphasize these correspondences.

To demonstrate the influence of 'environmental conditions' on the child's mental and educational development, Dr. Douglas compares, class by class, the changes in the children's average test-scores. It is, however, hardly correct to consider these averages as representing 'measured intelligence.' As we have seen, three of the four tests both at 8 and 11 measure verbal attainments; and, when he speaks of a 'deterioration in ability,' that may merely mean that a child who was fairly good in the picture test failed when an arithmetic test was substituted.

Let us, however, look at the detailed figures. A table in the appendix gives the 'average test score' at 8 and 11 for children in each of the four 'social classes.' Children belonging to one or other section of the 'middle class' improve by about half a mark; those of the 'lower working class' show a relative decline of half a

mark. No formal test is applied to show that, with these large samples, the differences are statistically significant. And Dr. Douglas overlooks the fact that at age 11 the standard deviation is appreciably larger; thus, the increase in range would amply account for the differences observed. In other words, it would seem that the differences themselves are artefacts resulting partly from the change in scale, and partly from the fact that the tests applied at the age of 8 had a decidedly lower reliability.

Dr. Douglas next sub-classifies the children according to their 'level of ability,' i.e., according to their initial scores at 8, taking an interval of 5 marks as indicating a separate 'level.' He then tabulates the 'change in score' at every level and in every social class. The conclusion drawn is that "the middle class children improve their scores, and the working class children deteriorate, at each level of ability." A glance at the table shows that this is far too sweeping. At the level indicated by a score of '60 or over' both sections of the middle class children *decline* in ability, and at *all* the levels below 50 both sections of the working class children (with one trifling exception) *improve*. Only at two of the six levels (the two just above the average) does Dr. Douglas's twofold inference hold, and even these include one exception. However, when tests of this kind are repeated and results of this nature are observed on comparing the averages for different levels, the changes may well be due to the phenomena of regression and counter-regression, i.e., they too may be artefacts. Much the same changes are often observable when the comparison is reversed and the 'levels' are based on test-scores obtained with the *second* set of tests—a check Dr. Douglas has not thought of applying. But suppose we assume his inference to be correct: what would it prove? Not that the middle class children *as a whole* tend to show a greater improvement, while the working class children *as a whole* tend to deteriorate (that, as we have just seen, is unsupported by the averages given in the appendix), but merely that many of the moderately bright individuals in the middle class improve, while many of the *duller* children in the working class drop further behind.

The environmental factor on which Dr. Douglas rightly lays chief stress is 'parental encouragement.' "The middle class parents," he says, "take more interest in their children's progress": the fathers, as well as the mothers, visit the school more frequently, and they are more eager for their children to have a grammar school education. Similar parental differences are also found within each social class; and on re-classifying the children according to 'parental attitude,' we discover, it is said, that those whose parents show the greatest educational interest, not only make higher scores at both ages, but also make greater progress. That we might certainly expect with tests of attainment; but, we are told, it also appears with the tests of ability. However, even if we grant the validity of the so-called 'ability tests,' it still would not follow that the improvement in 'intelligence' is due to the way the children "are stimulated by their parents": for the parents with these "high educational aspirations" seem largely to have been those who had themselves enjoyed a better education—doubtless because (particularly in the case of "parents from a working class background") they had won scholarships to a 'secondary' school." In short, much of the children's rapid development might have been due to the higher intelligence they had inherited from their parents.

It is, however, a truism in statistical research to say that underlying causes cannot be inferred from statistical computations alone. What was needed was to select cases showing the largest discrepancies, and then to make an intensive study of each one. It is usually found that on entering school youngsters from middle class homes are far better at reading and in verbal attainments generally than those from the working classes; on the other hand, with the brighter children from the working classes good school teaching will usually succeed in abolishing, or at least reducing, the handicap by the time the children have reached the age of 11. That might account for some of the anomalies in Dr. Douglas's tables. Many of the largest discrepancies may have been due simply to the fact that during one or other of the test-periods the child was suffering from some minor physical or emotional disturbance. But there is no need to enumerate the wide variety of causal processes that may be operative in such cases—some, no doubt, similar to those which Dr. Douglas

has imagined, others (and perhaps the more frequent) of a very different nature. What I am most anxious to deprecate is the neglect of the child *as an individual*, which has become so common in statistical researches in the educational field when undertaken by an outside investigator. The inside investigator—an educational psychologist (for example) attached to the staff of the local authority—sees such issues from a different standpoint. If, in the course of his inquiries he discovers certain pupils unexpectedly improving and others falling off, he wants at once to know who they are—why Lizzie's I.Q. has apparently increased, and why Tommy's marks in arithmetic are so much poorer than before. What we really need, therefore, when we come to the interpretation of Dr. Douglas's data, is not mere armchair theorizing about "the handicaps of the working classes," but a set of specimen case-histories to reveal what precisely are the commoner causes actually at work in this type of case or that.

The second of the two main problems, mentioned at the outset by Professor Glass, is the way in which environmental factors affect the child's allocation to the grammar school. On comparing data for pupils classified according to his four social classes, Dr. Douglas found that "54 per cent. of the upper middle class children, but only 11 per cent. of the lower manual working class children, go to grammar schools." This "waste of talent," he says, "may occur either through the effects of an adverse home environment or through lack of stimulation in early life." As "evidence that extreme poverty leads to a progressive deterioration in academic ability," he cites an investigation on English canal boat children, described in an American paper on 'Socio-economic Status and Intelligence.' The inquiry, which was carried out by one of H.M. Inspectors of Schools some forty years ago, and published by the Board of Education, was not, in fact, concerned with poverty; as a glance at the original report would have shown, many of the parents of the canal boat children were earning quite good wages. The object was to demonstrate how such children, leading peculiarly isolated lives and always on the move, were badly handicapped in tests of the Binet type: some had such low I.Qs. that they were in danger of being certified as mentally defective. Tested by the novel 'performance tests,' many revealed an ability quite up to, and even above, the general average. The effects of 'extreme poverty' (rarely encountered nowadays) have been fully discussed in books and reports on backward children—a literature to which Dr. Douglas nowhere refers. In the sample which he himself investigated there were, we are told, 7 per cent. living in circumstances which were economically or materially 'adverse': the chief reasons were "prolonged unemployment of the parents, mainly through illness, or break-up of the family owing to death, divorce, or separation." On the whole, however, during the eleven years covered by the survey, economic conditions steadily improved; hence today it is rather "the parent's attitude towards their children and towards education" that is chiefly to blame. Where the parents' interest is high, as many as 46 per cent. of the children are awarded places in a grammar school; where it is low, only 30 per cent.—appreciably more than their test-results would lead us to expect in the former case, slightly less in the latter.

Dr. Douglas ends his book with a chapter on the 'Wastage of Ability' in which he sums up his more important findings. "The evidence set out above," he says, "gives strong reasons for believing that much potential ability is wasted during the primary school years, and misdirected at the point of secondary selection." There is too "strong evidence that between 8 and 11 the ability and achievements of children is greatly influenced by their homes and schools—influenced moreover in a predictable way. During these three years at the primary school the middle class children pull ahead; and as a result the inequalities in the distribution of selective secondary school places are much greater when the eight-year test scores are used as the measure of ability." Accordingly, "taking the upper middle classes as the standard to be aimed at," he calculates that "1,005 children who ought to have got places failed to do so." And he ends by stating that "to meet this demand we should have to provide secondary schools for 47 per cent. of all children—a calculation," he repeats, "based on the eight-year tests."

How far is this final conclusion warranted? To begin with, let us note, that the attempt to select the brightest 47 per cent. would mean drawing the line of division

at the most difficult point. No method of selection can, of course, ever be perfect : the changes in development that occur during the pubertal years would alone be sufficient to prevent that. Let us suppose that its accuracy is indicated by a correlation of about 0.90—the highest we can reasonably hope for : then the proportional number of errors made when trying to select the highest 20 per cent. would be about 1 in 10 ; but, if we try to select the brightest 47 per cent., it will become roughly 1 in 4. But there are two far more questionable assumptions underlying Dr. Douglas's proposal. He assumes that these four simple tests, administered to youngsters of only 8 by inexperienced teachers, provide a better assessment of each child's suitability for a selective secondary school than the elaborate selection procedures carried out a few months before the actual transference—procedures based not only on tests far more carefully constructed and administered, but also on teachers' records and assessments, and usually on some supplementary procedure for doubtful or borderline cases. And he also, it would seem, assumes that the test-scores thus obtained yield a valid assessment of what he calls 'potential ability.' This, as we have seen, is quite unjustified ; indeed, Dr. Douglas himself, later on asserts that "no claim is made that these tests measure innate ability."

But, here as before, his figures could easily be interpreted in a very different fashion. The four-fold classification, based, as it is, partly on the educational attainments and secondary schooling of the parents and the social status of the grandparents, might be taken as yielding an approximate estimate for the innate and inheritable ability of the stock from which each child is drawn ; and the more rapid development of most of the upper middle class children might then be regarded as an indication of the way their innate ability has progressively matured. In the absence of more concrete information, a one-sided interpretation of this kind would, no doubt, be as indefensible as Dr. Douglas's. Both types of factor must be operative ; and there is certainly no reason to suppose that the preponderant influence must be environmental. The crux of the problem is the familiar fact that in a random sample of the population a favourable environment tends to go hand in hand with a favourable inheritance. The stock method of dealing with such complex issues is to keep first one factor and then the other as nearly constant as possible, e.g., to take first of all children of different heredity brought up in the same environment and then children of the same or similar heredity brought up in different environments. Failing that, the most plausible procedure is to calculate coefficients of correlation, and partial out the different components.

In view of the numerous studies already carried out, I venture to suggest that we no longer need any further attempts to demonstrate that the *general* environmental conditions associated with differences in social class influence children's *general* educational progress. What is now required are more specialized inquiries to show what particular elements in the environment influence what particular aspects of development in which particular types of child. And this means a far more elaborately planned investigation and a far more elaborate type of statistical analysis. With these more specific issues in view, future investigators will find it instructive to compare Dr. Douglas's methods and conclusions with the methods adopted by Professor Wiseman in his recent Manchester survey and the more cautiously balanced conclusions which he has reached (*Education and Environment*, 1964).

In a brief appendix, Dr. Simpson, of Rothamsted Experimental Station, has carried out an analysis of variance of the test-scores to show that the differences between the means of the children when classified according to (i) housing, (ii) size of family, (iii) academic record of school, and (iv) parental encouragement are significant in nearly every case. Yet once again it is necessary to repeat that numerical findings such as these are open to various interpretations. We all know that, generally speaking, the more intelligent parents tend to restrict their families and to take a greater interest in their children's education : they live in neighbourhoods where better housing conditions are to be found ; and the schools which their children attend inevitably achieve better academic records. Naturally, too, we should expect their children to obtain higher scores in tests of intelligence and attainment such as Dr. Douglas has used. Dr. Douglas, however, prefers to ascribe the results to environmental influence. "The primary schools with the best records,"

he says, "attract pupils from the middle classes" (*my italics*); "and good teaching in the primary schools will make up for any deficiencies." But no convincing evidence is offered to support this alternative method of interpretation.

Far more important than housing or size of family, one would have thought, must be the after-effects of early illness. The sample of school pupils which Dr. Douglas has studied is (as his previous reports have indicated) almost unique in the amount of information collected about the ante-natal conditions, early growth, and pre-school illnesses of the children. Wide differences were found; yet no attempt has apparently been made to relate these earlier findings to the later assessments for ability and attainments. However, we are promised a further volume on "the children's progress in the secondary system"; and it would be of special interest if these long-range effects could be discussed in full detail then.

The way in which I have seized upon the weaker points in Dr. Douglas's discussion scarcely does justice to his book as a whole. It includes many shorter chapters dealing with interesting points and containing much instructive information. His review of 'regional inequalities in selection,' together with the relevant tables, well deserve attention. If, therefore, I have over-emphasised the logical shortcomings in his arguments, that is because his 'national survey' has already been repeatedly cited in certain quarters as "a decisive demonstration of the overwhelming influence of the wide differences between environmental conditions of the privileged and underprivileged classes, and the deplorable wastage of ability that ensues as a consequence under our present ill-conceived plans for selection and allocation." Dr. Douglas's own conclusion is far more modest. "All that has been shown," he says in his final chapter, "is that a child's capacity is to a certain degree, dependent on the encouragement he gets from his parents, the sort of home he has, and the record of his school. Even if their influence is no greater than that suggested by the changes between 8 and 11, this represents a loss of ability which no selective examination can eliminate."

L. S. HEARNshaw (1964). *A Short History of British Psychology, 1840-1940*. London: Methuen, pp. xi+331, 35s.

By R. S. PETERS

This is the first history of British psychology to be written. It begins, somewhat arbitrarily in my view, with the psychology of Alexander Bain who, though he did no experiments but based his views, like Hume, explicitly on the observation of himself and others in ordinary circumstances of life was, nevertheless, "the first man who might claim to be first and foremost a psychologist." It ends with the second world war which, roughly speaking, marks the end of the period when British psychology developed more or less in an indigenous fashion before the tide of American theorizing washed over and submerged it.

This history was well worth writing and Professor Hearnshaw has managed to combine erudition with a level of exposition which is not too specialized for the general reader. The result is a mine of information which should be compulsory reading as part of the general education of all British psychology students.

Professor Hearnshaw is right to give some of the almost forgotten figures of the Victorian period their due. Most modern students have just about heard of Herbert Spencer, yet the name of G. H. Lewes, who was a more perceptive though less systematic thinker, awakens no response—unless, that is, the student is an admirer of George Eliot, which is improbable. They will similarly have heard of Galton; but how many have ever read him or have become imbued with the disciplined audacity of his approach to psychological phenomena? In Galton, one confronts what is of most value in the piece-meal, empirical tradition of British psychology. There was no fascination for all-embracing theories; no rigid adherence to a methodology. Instead there were concrete questions which intrigued him and brilliant techniques developed in trying to find answers to them.

The contribution of British thinkers to the development of neurology and neuropsychology is beyond dispute and Professor Hearnshaw shows considerable clarity and economy of exposition in putting the contributions of Hughlings Jackson, Sherrington, and Head, in their historical setting. He also does ample justice to the founders of comparative and social psychology such as Romanes, Lloyd Morgan, Hobhouse, and Graham Wallas. His treatment of changes in the philosophical climate is somewhat one-sided. He shows the pestiferous influence of German idealism on British thinkers in the latter part of the nineteenth century but says almost nothing about the long struggle during the twentieth century to rid British philosophy of this influence. There are one or two references to Russell but none at all to G. E. Moore. He does not see that Bradley's dictum "In England we have lived too long in the psychological attitude," is often quoted as a starting point of the "revolution in philosophy," because of its significance in the increasing awareness of what is characteristic of philosophical problems, and in the struggle by philosophers to put their own house in order. For the fact is that ever since the time of Locke epistemological and conceptual points had masqueraded in psychological guise. A remark such as Bradley's need not, therefore, when quoted, be indicative of hostility to psychology as an empirical science.

Professor Hearnshaw actually attributes psychology's slow academic growth during the first half of the twentieth century to such hostility on the part of philosophers who were at vantage points from which they could influence it. There is, regrettably, much in this accusation—especially in the case of Oxford and other universities where German idealism had taken root. But, on the other hand, I cannot refrain from saying that hostility to certain types of psychology is well-founded. No philosopher in his senses should be hostile to studies like those of Burt on juvenile delinquency or of Bartlett on remembering, or of Head on sensation. But hostility towards Behaviourism which was mainly just bad philosophy, or to Gestalt Psychology, which systematically confused psychological with conceptual questions, is surely more justifiable. These types of psychology are surely in a very different category from the more piece-meal empirical enquiries which have been so

characteristic of the indigenous British tradition in psychology which Professor Hearnshaw has so painstakingly catalogued.

The basic trouble, of course, is that in fields such as those of intelligence, action, thinking, perceiving, and emotion, it is very difficult to sort out what are empirical issues. A study of the work of Piaget, for instance, or of the James-Lange theory of emotion reveals this only too clearly. So there is bound to be a certain amount of sniping at the borders where philosophy and psychology meet. In the period about which Professor Hearnshaw has written, there was even less clarity than there is now about what are philosophical and what are psychological questions. So the conflict was, and to a certain extent, still is, situational rather than purely irrational or neurotic.

Professor Hearnshaw has a chapter on the 'systematic' psychology of Sully, Ward, and Stout in which this borderline area was systematically explored. Professor Hearnshaw's treatment of these thinkers would be much more interesting if he had raised more fundamental questions, like those raised by Mace, about the status of such enquiries. His treatment, on the other hand, of the other British systematizers, William McDougall and the London School, is lively because many of the issues raised are more concrete. He is, I think, a bit hard on McDougall. I myself would be inclined to be more well-disposed towards him because I think his main contribution to psychology was to defend a conceptual thesis about the importance of 'purpose' in the description and explanation of behaviour. Unfortunately, he also seems to have assumed that if the concept of 'purpose' is both irreducible and indispensable in the explanation of behaviour, then an organism must be born with a finite number of built-in purposes. In other words he translated an important conceptual thesis into a dubious genetic theory, as many others have done.

There is a very interesting chapter on the history of the departments of psychology in the various universities, in which Professor Hearnshaw systematically records what is usually a matter of academic gossip and folk-lore. Educational psychology and child study get very full treatment—perhaps much fuller than is warranted by its limited contribution in this country to psychological theory.

In brief, this is a very useful, informative, and competent history. It reads rather like a very long encyclopaedia article. By that I mean that the reader must not expect much in the way of historical speculation or acute discussion of theoretical issues raised by British psychologists. The history is rather like British psychology itself—competent, piece-meal, sensible, but at times, a trifle dull.

BOOK REVIEWS

BOOK REVIEW WRONGLY ASCRIBING AUTHORSHIP TO PROFESSOR R. A. C. OLIVER

It is very much regretted that, because of an editorial error, the publication *The Marking of Scripts in Advanced Level History*, which was reviewed in the last number of this *Journal* by Mr. A. E. G. Pilliner (February, 1965, Vol. XXXV, Part I, pp. 110-1), was wrongly ascribed to the authorship of Professor R. A. C. Oliver. The error is the more regretted in that the publication was very unfavourably reviewed and reflects adversely upon the competence of the researcher who carried out the work. The name of the author of the publication has not been published. Professor Oliver had no part whatever in writing it. Subscribers to the *Journal* are requested to note this correction.

AUSTRALIAN COUNCIL FOR EDUCATIONAL RESEARCH (1964). *Review of Education in Australia, 1955-62*. Melbourne: A.C.E.R., pp. 436, £A5.

The key to this book is given on p. 60, "In fact, the word 'education' does not appear in that part of the Commonwealth Constitution which deals with the powers of Parliament." Indeed, the Federal aspects of Education seemed to be those mainly concerned with universities, with schemes for rehabilitating ex-servicement, with immigrant and armed forces education, and with administering overseas scholarships. In essence, therefore, this book concisely reviews autonomous State education systems by taking as a chapter heading an aspect of education such as the primary field and stratifying it into State sections. It deals with the pre-school, primary, secondary, technical and university facilities in each State as well as discussing administration, finance, teaching staff and special education services. The Review provides a valuable fund of information on the education systems of each state and makes considerable use of tabulation to condense the information.

The major stresses in Australian education were stated as the increased demand for secondary and university education and the re-organisation of such facilities to meet this demand, together with the diversification of technical education and the universal shortage of specialist teachers. It seems that Australia has its share of these world-wide problems.

There were a few points on which amplification would have been merited. One might have wished for, example, that the important educational contributions made by the Protestant Independent System were discussed to balance the statement on Catholic Education and the comprehensive review of the States Secular Systems, which provide the bulk of the material in the book. Further, a bibliography or reference list at the end of the book would have been invaluable.

The educational stimulation which was felt to have resulted from the travel of teachers and administrators should perhaps be borne in mind by other educational systems. From other sources one reads of Australia's implementation of this view with schemes in 1964 and 1965 for an inter-state exchange of teachers. Such an interchange had previously been noticeably absent in some States which did not accept at par teaching experience gained overseas, inter-state, or even in the independent school systems of their own State.

The value of the bursaries awarded to student teachers seemed to vary widely, but the better grants are considerably above those in the U.K. For example, in New South Wales, a married man with one child could be receiving the equivalent of £575 p.a. (plus fees). However, the recipients of such bursaries were subsequently bonded to their State education departments, some of which were very reluctant to train non-bursar secular Australians. No married women seemed to have been trained in Australia in this period and in some States women teachers were required to resign on marriage or only considered for temporary posts, thus accentuating any teacher shortage problem in Australia. Perhaps this situation is also under 'review,' judging from some of the more recent Departmental Statements of 1964.

An interesting trend in secondary education which might send a tremor through the subject purists in England and is the decision by one State in 1961 "that the subjects of Physics, Chemistry, Biology, Physiology, Hygiene and General Science would be replaced by one subject known as Science . . ."

Any major review is, by nature, a fact-fat survey. It is rarely light reading. This book is not heavy going and has not sacrificed information or expanded its pages to be readable. It is a valuable addition to the library of any administrator or specialist in comparative education who would, however, be advised to acquaint himself with the few thin patches in the book, and to realise that he will himself have to make most of the inter-state comparisons.

K. B. START.

SECONDARY SCHOOL EXAMINATIONS COUNCIL (1964). *Examinations Bulletin No. 3—The Certificate of Secondary Education: An Introduction to Some Techniques of Examining*. H.M.S.O., pp. 69, 6s. 0d.

VERNON, P. E. (1964). SECONDARY SCHOOL EXAMINATIONS COUNCIL. *Examinations Bulletin No. 4—The Certificate of Secondary Education: An introduction to Objective-type Examinations*. H.M.S.O., pp. iv+20, 2s. 3d.

Examinations Bulletin No. 3.

This suffers by comparison with Bulletin No. 4 reviewed below. The objectives are certainly modest: "... to make available a brief account of the present state of knowledge about those techniques of examining which seem likely to contribute to the fairness and efficiency of the new examinations." These are achieved, but one could have wished that the real needs of the teachers responsible for the C.S.E. examinations could have been more fully appreciated. Some attempt has been made to make the necessary techniques understandable to the reader and to separate the statistical background from the narrative but the first half of the Bulletin still needs the kind of concentrated attention which few teachers will be prepared to give.

The aims as stated are admirable: "To have available a summary of those techniques, statistical and other, which research and experience have shown to be helpful in improving the efficiency of examinations, and in helping teachers and examiners to co-ordinate their judgments." Unfortunately, the result is a document which could easily encourage teachers to believe that these techniques are beyond them. Here is an attempt to achieve perfection from the start. The authors would have been better advised to seek more modest targets for their teacher audience to aim at. The intelligent but uninformed reader is, perhaps, willing to struggle through a reasoned exposition but he will soon give up if there are too many missing steps to be taken on trust. "... whether or not the paper had discriminated over the whole range of candidates by spreading them out on both sides of the mean. There would be reason to believe this to be so if in a scale with a range of 100 marks, there was a standard deviation of about 15. (A more detailed discussion of the standard deviation is to be found in Part II)." However, the promised discussion produces no further explanation for this statement.

The treatment of validity may well suggest standards impossible of achievement by the amateur examiner: "An inspection of an examination paper together with the marking scheme will provide a competent judge . . . with some idea whether the paper is suitable for its purpose: that is whether it is valid or not. But this may prove to be rather superficial estimate of validity unless care is taken to set the framework within which the judgment is to be made." And again—"If the two papers have been measuring valid outcomes of the course of study, there should be a positive correlation between them. This correlation . . . cannot, however, be expected to be very high. It will be reduced by the sampling of the content and objectives which was undertaken at the stage of item construction, and by the difficulty of judging whether any particular question does, in fact, measure a valid objective."

"An attainment examination should, nevertheless, possess a degree of predictive validity."

"... teachers' estimates are based on knowledge of only a small sample—the reliability of such estimates may be lower (or higher) than the reliability of the written paper. This difference in reliability will affect the correlation between them and thus affect the estimate of validity itself."

The discussion of the problems of validity is rounded off with a warning "... technical problems abound in estimating validity. It is an area in which even the expert is wise to move very cautiously." The average teacher must inevitably ask himself how much of this discussion really helped him to meet the practical problems in the field.

In the light of the very full discussion it is perhaps a pity that item analysis, except in the construction of an objective type of examination paper, is regarded as "... rather imprecise and of only limited value."

There is an introduction in Bulletin No. 3 to the question of Moderation between schools. Although this will be dealt with in detail in Bulletin No. 5, which will soon follow the publication of No. 4, enough is said to show the kind of central direction which most Regional Boards will attempt to follow. Some of the features of this moderation are new to teachers and these are regarded as perhaps the most controversial in discussions taking place on the means of conducting the Certificate of Secondary Education.

The advice given to the would-be moderator is not at all clear and seems at times to be over-subservient to the school's own assessments.

"... it is useful to make the assumption that the grades proposed by the school are correct, and will not be challenged if there is a reasonable level of agreement between them and the grades for the school which the moderator proposes."

"The process of moderation should allow for differences in the content and style of work done in different schools. This is achieved by requiring only a reasonable degree of agreement between the moderator and the school."

In the second half of the Bulletin it is suggested that two criteria should be applied to test the 'level' and 'degree' of agreement. It becomes clear that the criteria are based only on a difference of means (or 'level') and on the degree of correlation (or 'degree' or 'conformity'). Bulletin No. 5, it is understood, will acknowledge the inadequacy of relying only on two criteria and will add a third criterion, namely, one which is based on a difference of spread of the two arrays of grades.

Finally, the Bulletin adds a cautionary note to explain: "The elementary statistics which have been introduced in the second part of this Bulletin have been dealt with very simply. A number of problems, however, have not been covered, particularly those concerned with the confidence placed on any statistic..."

For the experienced statistician these ideas are elementary, yet their difficulty of understanding by the teacher in the classroom has not yet been fully appreciated. Many determined and conscientious moderators-to-be will work through the examples provided with great difficulty. In this connection, it is unfortunate that there are serious errors in the formula for the correlation and in its calculation (page 45). It is understood that these will be corrected in the next edition.

It is evident that a great deal of thought and consultation has gone into this Bulletin. The authors are to be congratulated on a useful work of reference which provides an up-to-date survey of most of the problems to be faced. What is now needed is a down-to-earth primer which will give practical and positive guidance for the army of moderators of all subjects now being appointed by Regional Boards.

Examinations Bulletin No. 4.

This little book serves its intended purpose admirably as an introduction for teachers in one capacity or another, with the developing Certificate of Secondary Education. A great deal more still needs to be done if prejudice and opposition to objective type examinations is to be successfully overcome. Numbers of teachers have been working devotedly throughout the country to ensure that the new examination works smoothly from the start. The biggest task of all, however, lies ahead and this is the persuasion of the teaching profession as a whole to accept the ideas which are being promulgated from the centre. At times these efforts may appear

to the converted to be too conciliatory, but there is no doubt that conviction will only come with experience. Vernon emphasises this feeling by comparison with America . . . "In contrast, in the United States of America, objective examining is more commonly used than conventional throughout secondary schools and colleges, both for selecting students and for grading their aptitudes and achievements." . . . "Why have British educationists been so reluctant to follow suit? . . . largely the reason lies in the conservation of British secondary school and university teachers, and their suspicions of this type of examining."

The Bulletin discusses the balance between objectivity and subjectivity achieved in the use of both objective and conventional type examinations and emphasises that these differences cannot be expressed in terms of black and white, but in various shades of grey. The author goes on to consider in turn some of the varied arguments which have been advanced in the past, the ease and speediness of objective questions, their influence on teaching and learning and the effect of guessing. All these are presented in a fair and useful "for and against" table which could usefully provide an informed background for teacher discussion. Vernon presents a timely reminder of the valuable contributions of Professor B. S. Bloom, of the University of Chicago, in his classification of educational 'objectives': knowledge, comprehension, application, analysis, synthesis and evaluation.

It is particularly relevant that the Bulletin concludes with a glance to the future and emphasises the role of the objective examination as a complement to the conventional essay type examination. Once again it is the shortage of competent advisers in this field which will do most to inhibit progress and Vernon suggests . . . "it would be worthwhile arranging for quite a number of potential examiners to obtain a month or so's training in the United States." This is a suggestion which should not go unnoticed in a publication of the Schools Council.

Finally, there is the suggestion that the forward looking Certificate of Secondary Education may well, in time, bring about other reforms: "If the objective examination works well at C.S.E. level, it is likely to begin to appeal more strongly to G.C.E. examiners, and even in time to find uses at A level or in tertiary education, whenever the numbers of examiners are large."

NORMAN FRANCE.

DE CECCO, JOHN P. (1964). *Educational Technology*. New York: Holt, Rinehart and Winston, pp. xiv+479, \$4.25.

Programmed Learning is now beginning what might be called the second Five-Year Plan. This collection of readings on the subject is timely. There are seven original papers and thirty-five reprints, some in shortened form. All have been grouped into ten chapters by the Editor and he has written introductory material for every section and for a number of individual papers.

Big names in the field are represented, but at rather a low level of exposition in a number of cases, perhaps because of the needs of that journal for which they were originally written. There is one paper from Britain. Inevitably, there is repetition and also advanced ideas and hackneyed ones are inter-mingled. This is an increasingly irritating feature of books on programmed learning. Pressey is the prophet, Skinner the proselyte, Crowder the heretic and Finn the visionary.

The book begins badly with four papers at an elementary level giving practical advice for programming. The necessary simplification results in uncritical writing and in blinding flashes of the obvious, such as the fact that programmes of only 50-60 frames are not worthy of research effort. The Editor does advise the cognoscenti to omit this section.

Initially Programmed Learning was mainly concerned with techniques, and its major characteristics have not yet been listed, far less examined critically. Most of this book is taken up with theoretical considerations of learning and methodology, much of it based on recent research projects.

Stimulus-Response theory is related to task analysis and sequencing, going beyond Guthrie and Hull with their reinforcement and distribution of practice. Thorndike's 'frequency' is criticised as being of little value independent of reward.

Laboratory studies of verbal learning are related to school practice, and a curriculum model, or instructional system, is postulated, leading to an educational technology. Lewin's precept that "nothing is as practical as a good theory," is accepted, the end product being the behaviour of the learner. While the various contributors have differing theoretical positions, yet the balance would seem to be in favour of a S-R contiguity theory.

Subsequent chapters examine or re-examine basic beliefs of learning theory and of measurement technology. Thus, classical conditioning, transfer of training, distribution of practice, I.Q., as an achievement predictor, mental discipline, motivation, knowledge of results are reviewed and discussed at length.

Dewey becomes a link which carries the theme over to educational philosophy from psychology and an attempt is made to define the objectives of education in behavioural terms. Without this it is not possible to define the limits of any programme.

The last three chapters endeavour to evaluate the whole field of programming, its use in school and its future development. Important features are that methods of teaching may be considered afresh, and it would seem that the research worker and the class teacher may be brought into a much needed closer contact.

J. G. MORRIS.

FLUGEL, J. C., and WEST, D. J. (1964). *A Hundred Years of Psychology*. London: Duckworth, pp. 394, 30s.

Flugel's *Hundred Years of Psychology* was a classic of its kind. It lacked, perhaps, the erudition of Volume III of G. S. Brett's monumental *History of Psychology*; but it also lacked Brett's meandering style. On the other hand, it was not so narrowly specialized as Boring's histories of the same period. It was an eminently civilized book combining the virtues of accuracy, breadth of interest, and clarity of organizational structure. Like all histories, of course, it was highly selective; but the selection was made with a nice sense of historical importance and of general human interest. In brief, it was the work of one of the most cultured psychologists of the period between the two world wars.

The first edition was published in 1933. In 1947, a revised edition was published in which Flugel included a supplementary review of developments from 1933 to 1947. In the third edition (1964) Dr. West has left the first edition material up to 1933 intact, but has revised the supplementary material of the second edition and has included it in a survey of developments from 1933 to 1963, which comprises Part V of the third edition.

Part V continues Flugel's catholicity of interest but it lacks both his sense of structure and of historical development, and his human touch. It may well be that we are too close to the events of the last thirty years to pick out the main lines of historical development; but surely, there are such trends—not just one damn thing after another, which is the general impression given by Part V. Surely sub-headings would have helped, or breaking the material up into short chapters, as Flugel did in his original edition. Indeed, the headings of Flugel's own Part IV, if continued into Part V, might have helped to organize the material which has been so loosely strung together by Dr. West.

The exposition of Part V is competent enough and there is plenty of space given to work done by psychologists on matters of general human concern. Indeed, the emphasis is more on the application of psychology to human problems than on the various attempts to develop it as a theoretical science. But the presentation of the material is very dead pan, unenlivened by delightful flashes such as Flugel indulged in when talking, for instance, of the work of Galton. Psychology, of course, is now much more drab and professionalized than it used to be in those days. But have all its characters been colourless? If such a history is to be absorbing one, it must have a pattern illuminated either by an imaginative recreation of the problems of the past or by some asides about the people who dealt with them. Flugel's original work derived its distinctiveness from this combination. Dr. West has only gone through the motions of bringing the work up to date; for his final chapter lacks the spirit which gave life to the original.

R. S. PETERS.

BRADFORD, L. P., GIBB, J. R., and BENNE, K. D. (Eds.) (1964). *T-Group Theory and Laboratory Method*. London: John Wiley and Sons, pp. xii and 498, 75s.

This book is described as "the story of an innovation in education—the T (Training) Group." It is an interesting story which is worth telling. The term 'laboratory' is used to refer to a training laboratory, in which individuals participate for one or two weeks in situations intended to promote learning. The objectives of a training laboratory include learning about oneself and one's impact on others, about inter-personal and inter-group relations. One assumption behind the learning situations which are developed is usually that it is not possible to completely attain these objectives by an intellectual process; the emphasis is commonly on emotional experiences which are difficult to verbalize. A prominent part of most training laboratories is the T-group, which may be roughly described as an originally unstructured group whose task is to study its own activities and experiences. A T-group usually includes one or more trainers who assist the learning process by means of infrequent interventions.

This book comprises chapters by twelve American behavioural scientists about behaviour at training laboratories. All the writers have, at some time, been associated with the National Training Laboratories at Bethel, Maine, but there is, nevertheless, an interesting diversity of opinions. (This diversity would probably have been wider if contributions from English T-group trainers had been included.) But there are many points on which agreement is now fairly general, and the editors have failed to prevent several contributors from covering the same ground. For example, the general aims of T-group training are repetitively discussed in far too many chapters. In general, it is likely that the book's impact would have been greater if it had been considerably shortened. This difficulty is one of meeting the different requirements of people who have experienced T-group training and those of readers to whom the ideas are new. The former group will be familiar with many of the details presented, whilst the latter will often not be able to grasp what is involved. T-group members often have tremendous difficulty in describing to an outsider what is happening. Several chapters attempt to give an impression of T-group behaviour by means of transcripts, and these will prove useful to readers unfamiliar with such groups. But these readers will find large parts of the book very heavy going, since the significance of many T-group behaviours is hard to fathom without experience of them. And those readers *with* T-group experience will probably find it hard to maintain their interest during the frequently repetitive passages.

Despite these failings, parts of this book are likely to be helpful to those who are at all interested in the interactions occurring in a learning situation. Recent progress in this field has been exciting, and generalizations and hypotheses abound. Perhaps a future book will report tests of some of the hypotheses.

PETER B. WARR.

COLLINS, BARRY E., and GUETZKOW, HAROLD (1964). *A Social Psychology of Group Processes for Decision-Making*. New York: John Wiley and Sons, pp. x+254., 47s.

Those who view with dismay the appearance of yet another book on group processes, can take heart; this book is well worth reading. It also has the merit of taking the reader into the author's confidence by explaining, in an exemplary first chapter, exactly how the book came to be written, the criteria used in selecting the bibliography, the approach taken to contradictory evidence and the content and lay-out of subsequent chapters. Readers are invited to make public their comments and criticism on the assumption that they will inevitably possess resources not utilised by the authors and, therefore, have their own contribution to make to an inductive theory of face-to-face groups.

Many of one's queries and objections are, in fact, forestalled by Chapter One. Why, for example, is the final chapter on leadership different in form from the rest and why does it appear rather like an afterthought? The answer is that this apparently is what it is; the authors, having decided not to use 'leadership' as a conceptual theme on the grounds that relevant studies were included under the

headings, found they had excluded two interesting streams of research on leadership. They, therefore, added Chapter Eleven to make good the omission. Such refreshing honesty compels a lenient view of what are really only minor defects in this admirable book.

It is, however, not a book for beginners, who would do much better to start with, say, Josephine Klein's *Working with Groups* and then turn to Collins and Guetzkow for more detailed and advanced study. The material is highly condensed, but it is well organised, following the present convention of giving empirically tested hypotheses as a series of numbered propositions, together with a summary of each chapter. Supplementary evidence is presented in foot-notes and in optional sections of the text set in smaller type than the rest. Whether such 'processing' is to everyone's taste, is another matter, but no doubt, many will find it helpful.

Teachers and others concerned with the performance of people in groups will be particularly interested in the first four chapters and the review of findings related to the problem of whether people work better collectively or alone. The remainder of the book will fascinate all those with experience of committees and conferences, especially the chapters dealing with sources of power and influence in decision-making groups.

SYLVIA SHIMMIN.

ADLER, SOL. (1964). *The Non-Verbal Child*. U.S.A.: Charles C. Thomas, pp. xii+163, \$6.75.

Dr. Adler sets out to provide a simplified clinical guide to the problems of the 'non-verbal' child, mainly for what are called 'habilitationists.' In general, simplified guides to complex problems and procedures are rarely satisfactory and may be dangerous; this book is no exception. One of its defects is that we are not told exactly what is meant by a 'non-verbal' child or by 'significant retardation in language development.' There is little or no discussion of the various forms of language or speech disorder, and the author deals with the more pathological causes of 'non-verbalism' without giving the reader the necessary background knowledge and without reference to other factors which may affect a child's language development. Nor is it clear what professional roles the 'relatively unsophisticated student or habilitationist' would fulfil, though advice is given (mainly in one fairly short chapter) on audiometric testing, testing for sensory dysfunctioning, assessing language development, using psychometric procedures, testing the speech mechanism for symptoms of neuro-motor impairment and on carrying out a gross neurological examination. Although the book is not a long one, it is, in fact, difficult to see the wood for the trees.

The book so much suggests that the understanding of highly complex skills, developed by numerous specialists, can be conveyed simply and briefly to an amateur habilitationist trying to integrate them all that it cannot be recommended. This is a pity, as the explanations of brain mechanisms and communicative behaviour are not without interest, although some of the discussions are terminated prematurely and left in the air. Among other topics, Dr. Adler also deals with the question of making a differential diagnosis between the four basic causes of non-verbalism in the young child (brain injury, mental subnormality, emotional illness and deafness), and with the habilitation of the non-verbal child. As an appendix, we are given examples of sample forms used to collect personal data at a Speech and Hearing Clinic, as well as a detailed case-history and various reports on a nine-year-old minimally brain-injured child.

Although the style is generally clear, there is occasional slovenliness of expression and some of the statements made are misleading. For example, the observation that "any non-verbal child is also a mentally subnormal child," is likely to be misunderstood, even in its context. The author, too, frequently refers to 'the child' in discussions of testing techniques without mentioning which age range is being considered. This seems a particularly serious omission since Dr. Adler is much concerned with the diagnosis and treatment of children as young as 2½ years, to whom most of the tests dealt with would not apply.

The book is attractively printed and produced, and there are some useful suggestions for further reading.

MAURICE CHAZAN.

PUBLICATIONS RECEIVED

The mention of a book in this list neither implies, nor precludes, a later review.

- MAURICE REUCHLIN (1964). *Pupil Guidance Facts and Problems*. Strasbourg: Council for Cultural Co-operation of the Council of Europe, pp. 361, 10s.
- G. N. A. VESEY (1965). *The Embodied Mind*. London: George Allen and Unwin, pp. 113, 20s. net.
- D. W. WINNICOT (1965). *The Family and Individual Development*. London: Tavistock Publications, pp. vii+181, 30s. net.
- HENRI PIERON, ALFRED FESSARD and PAUL FRAISSE (1964). *L'Annee Psychologique*. Paris: Presses Universitaires De France, pp. xii+349, no price.
- ALFRED ADLER (1965). *Superiority and Social Interest*. London: Routledge and Kegan Paul, Ltd., pp. xix+432, 45s.
- A. J. MARSHALL, and C. N. ARMSTRONG (1964). *Intersexuality in Vertebrates, Including Man*. London: Academic Press, Inc., Ltd., pp. xi+479, 90s.
- GEORGE MANDLER and WILLIAM KESSEN (1964). *The Language of Psychology*. New York: John Wiley, pp. xviii+301, 15s.
- LEE, J., CRONBACH and GOLDINE C. GLEISER (1965). *Psychological Tests and Personnel Decisions*. Second edition. U.S.A.: University of Illinois Press/Urbana, pp. viii+347, \$7.95.
- RICHARD C. ANDERSON, and DAVID P. AUSBEL (1965). *Readings in the Psychology of Cognition*. New York: Holt, Rinehart and Winston, pp. vii+690, \$7.50.
- WILLIAM DEMBER (1965). *Visual Perception: The Nineteenth Century*. London: John Wiley, pp. xii+222, 38s. cloth, 23s. paper.
- W. R. NIBLETT (1965). *How and Why Do We Learn*. London: Faber and Faber, pp. 196, 21s.
- J. D. CARTHY and F. J. EBLING (1965). *The Natural History of Aggression*. London: Academic Press, Inc., pp. 159, 30s.
- EILEEN YOUNGHUSBAND (1965). *Social Work With Families*. London: George Allen and Unwin, Ltd., pp. 187, 25s. net.
- REGINALD D. ARCHAMBAULT (1965). *Philosophical Analysis and Education*. London: Routledge and Kegan Paul, pp. xii+212, 28s. net.
- GERMAINE DE MONTMOLLIN, ROGER LAMBERT, ROBERT PAGES, CLAUDE FLAMENT, et JEAN MAISONNEUVE (1965). *Traite de Psychologie Experimentale*. Paris: Presses Universitaires de France, pp. 272, no price.
- PETER M. BLAU (1964). *Exchange and Power in Social Life*. London: John Wiley and Sons, pp. xxiii+352, 59s.
- JACK A. VERNON (1965). *Inside the Black Room*. London: Souvenir Press, Ltd., pp. xviii+203, 21s.
- W. SLUCKIN (1965). *Imprinting and Early Learning*. London: Methuen and Co., pp. x+147, 25s.
- DR. A. KRIEKEMANS (1965). *Geschiedenis Van De Kinderpsychologie*. Uitgeverij Lannoo, Tiel/Den Haag, pp. 212, 145 Fr.
- MARGARET PHILLIPS (1965). *Small Social Groups in England*. London: Methuen and Co., pp. 317, 25s. net.
- HAROLD F. SEARLES (1965). *Collected Papers on Schizophrenia and Related Subjects*. The Hogarth Press and the Institute of Psycho-Analysis, pp. 797, 5 gns. net.
- EDITH JACOBSON (1965). *The Self and the Object World*. London: The Hogarth Press, Ltd., pp. xiii+250, 35s. net.
- THE YEARBOOK COMMITTEE AND ASSOCIATED CONTRIBUTORS (1965). *Theories of Learning and Instruction*. London: The University of Chicago Press, pp. xi+430+vi, 41s.
- DELWYN G. SCHUBERT (1964). *A Dictionary of Terms and Concepts in Reading*. U.S.A.: Charles C. Thomas, pp. xiii+297, \$6.00.

- H. S. N. MCFARLAND (1965). *Psychology and Teaching*. London: Harrap and Co., pp. 250, 12s. 6d.
- CROCKER AND CARMADA (1964). *Readings in Rhetoric*. U.S.A.: Charles C. Thomas, pp. xiii+599, \$16.50.
- EDRITA FRIED (1964). *Artistic Productivity and Mental Health*. U.S.A.: Charles C. Thomas, pp. xiv+177, \$6.50.
- MARIA MONTESSORI (1965). *Dr. Montessori's Own Handbook*. Massachusetts: Robert Bentley, pp. viii+121, \$5.00.
- A. M. WINCHESTER (1965). *Heredity: An Introduction to Genetics*. London: Harrap, pp. iv+269, 18s. net.
- ROCH DUVAL (1964). *Adolescence D'Aujourd'Hui*. Quebec: Les Presses de L'Université Laval, pp. 191, \$3.00.
- ERROL, E. HARRIS (1965). *The Foundations of Metaphysics in Science*. London: George Allen and Unwin, pp. 512, 63s.
- DOUGLAS P. CROWNE and DAVID MARLOWE (1965). *The Approval Motive. Studies in Evaluative Dependence*. London: John Wiley, pp. xii+233, 57s.
- R. L. KAHN, D. M. WOLFE, R. P. QUINN, J. DIEDRICK SNOEK, R. A. ROSENTHAL (1964). *Organisational Stress: Studies in Role Conflict and Ambiguity*. London: John Wiley and Sons, pp. xii+470, 60s.
- J. M. MANDLER and G. MANDLER (1964). *Thinking: From Association to Gestalt*. London: John Wiley, pp. x+300, 38s., cloth, 23s. paper.
- H. WEITZ (1964). *Behaviour Change Through Guidance*. London: John Wiley, pp. xii+225, 49s.
- A. ZALEZNIK and D. MOMENT (1964). *Casebook on Interpersonal Behavior in Organizations*. London: John Wiley, pp. xiii+587, 60s.
- ABRAHAM ZALEZNIK and DAVID MOMENT (1964). *Casebook on Interpersonal Behavior in Organizations*. New York: John Wiley and Sons, pp. xiii+487, 60s.
- HENRY WEITZ (1964). *Behavior Change Through Guidance*. New York: John Wiley and Sons, pp. xii+225, 49s.
- HERBERT ROTH (1964). *A Bibliography of New Zealand Education*. Wellington: New Zealand Council for Educational Research, pp. v+234, 22s. 6d.
- SOL ADLER (1964). *The Non-Verbal Child*. U.S.A.: Charles C. Thomas, pp. xii+163, \$6.75.
- DERICK UNWIN (1964). *Kinematics*. London: Methuen and Co., pp. 63, 4s. 6d.
- JOHN E. WISE (1964). *The History of Education*. New York: Sheed and Ward, pp. xiv+494, \$5.00.
- SUSAN MEYER MARKLE (1964). *Good Frames and Bad: A Grammar of Frame Writing*. London: John Wiley, pp. xiv+278, 35s.
- ANDREW R. BAGGALEY (1964). *Intermediate Correlational Methods*. London: John Wiley, pp. 211, 44s.
- G. W. FAIRWEATHER (1964). *Social Psychology in Treating Mental Illness: An Experimental Approach*. London: Wiley and Sons, pp. xii+300, 60s.
- MAUD MANNONI (1965). *Le Premier Rendez-vous Avec le Psychanalyste*. Geneve: Editions Gonthier, pp. 185.
- NORMAN K. HENERSON (1965). *The Education of Handicapped Children—Recent Trends and Research*. London: Oxford University Press, pp. viii+46, 4s. 6d.
- I. MACFARLANE SMITH (1964). *Spatial Ability*. London: University of London Press, Ltd., pp. 408, 35s.
- SIR JOHN FULTON (1964). *Experiment in Higher Education*. London: Tavistock Publications, pp. 16, 3s. 6d.
- JOE PARK (1964). *Bertrand Russell on Education*. London: George Allen and Unwin, Ltd., pp. 147, 25s.
- D. E. ADLAND (1964). *The Group Approach to Drama*. Book 1. London: Longmans, pp. 90, 4s.
- D. E. ADLAND (1964). *The Group Approach to Drama*. Book 2. London: Longmans, pp. 120, 4s. 6d.
- D. E. ADLAND (1964). *Group Drama*. London: Longmans, pp. vi+145, 12s. 6d.

- COUNCIL OF THE TAVISTOCK INSTITUTE OF HUMAN RELATIONS (1964). *Social Research and a National Policy for Science*. London : Tavistock Publications, pp. 44, 7s.
- MICHAEL BURN (1964). *Mr. Lyward's Answer*. London : Hamish Hamilton, Ltd., pp. 288, 12s. 6d.
- EDWARD ARTHUR TOWNSEND and PAUL J. BURKE (1964). *Statistics for the Classroom Teacher*. New York : Macmillan, pp. 76, no price given.
- FINLEY CARPENTER and EUGENE E. HADDAN (1964). *Systematic Application of Psychology to Education*. New York : Macmillan, pp. xi+270, 22s.
- WILLIAM S. RAY (1964). *The Science of Psychology*. New York : Macmillan, pp. vi+575, no price given.
- JOSEPHINE KLEIN (1965). *Samples from English Cultures*. Volume I. London : Routledge and Kegan Paul, Ltd., pp. xii+435+13, 50s.
- JOSEPHINE KLEIN (1965). *Samples from English Cultures*. Volume II. London : Routledge and Kegan Paul, Ltd., pp. xi+334, 35s.

E. & S. LIVINGSTONE, LTD.

THE DEVELOPMENT OF THE INFANT AND YOUNG CHILD Normal and Abnormal

Second Edition. By Professor R. S. Illingworth.
370 pages 126 illustrations 30s.

THE NATURAL DEVELOPMENT OF THE CHILD

Fourth Edition. By Dr. Agatha H. Bowley.
224 pages 84 illustrations 10s. 6d.

CHILD CARE

By Dr. Agatha H. Bowley.
224 pages 10s. 6d.

THE YOUNG HANDICAPPED CHILD

By Dr. Agatha H. Bowley.
136 pages 36 illustrations 10s. 6d.

CEREBRAL PALSY. The Pioneer Years of Occupational Therapy in Scotland

By A. Katharine Levin.
132 pages 104 illustrations 30s.

TEVIOT PLACE, EDINBURGH

THE BIRMINGHAM PRINTERS LTD PRINTERS & BOOKBINDERS

*High Class
Technical
Publications*

42-44 HILL STREET
BIRMINGHAM 5
TELEPHONE : MIDLAND 0196-7

FACTORS ASSOCIATED WITH MALADJUSTMENT IN
EDUCATIONALLY SUBNORMAL CHILDREN

BY MAURICE CHAZAN

(Department of Education, University College of Swansea)

SUMMARY. This article discusses the findings of the second stage of an enquiry into the incidence, nature and aetiology of maladjustment among children in special schools for the educationally subnormal in South Wales. An intensive study was made of the thirty 'most maladjusted' E.S.N. children in the larger sample studied in the first stage, as well as of a control group of the thirty 'best adjusted' E.S.N. children, matched with the maladjusted children for age, sex and school.

Significantly more of the maladjusted children (i) showed some physical weakness or defect, (ii) were subjected to adverse psychological pressures and unsatisfactory discipline at home, related to parental instability, and (iii) had had interrupted or incomplete relationships with their parents. There was little difference between the two groups in respect of the incidence of adverse congenital factors, difficulties in early development and poor material conditions at home. Significantly fewer of the maladjusted children had a positive relationship with their father. Maladjustment was, in many cases, associated with a lack of progress in the basic subjects at the special school.

Maladjustment of the 'withdrawn' type tended to be associated with physical deficiency in the child and positive but weak parental discipline; aggressive behaviour was often linked with insecurity at home and hostile parental attitudes.

These findings emphasize the need for comprehensive guidance services for the parents of educationally subnormal children.

I.—INTRODUCTION.

In a previous article (Chazan, 1964) the findings of the first stage of an enquiry into the incidence, nature and aetiology of maladjustment in special schools for E.S.N. children in South Wales were discussed. This article describes the second stage of the investigation, in which an intensive study was made of the thirty 'most maladjusted' children in the larger sample, together with a control group of the thirty 'least maladjusted' children, in order to examine the factors associated with maladjustment in E.S.N. children. The small sample of children studied permitted an approach in some depth, but it is realised that the results of the enquiry must be interpreted cautiously and need to be confirmed by means of larger-scale and longer-term studies.

Although there is a considerable literature on causative factors associated with maladjustment in childhood, relatively little has been written about factors associated with maladjustment in the E.S.N. child. Stott (1957, 1959) has suggested that congenital factors, especially stress during pregnancy, are particularly important as they are likely to result in the multiple impairment of mental retardation, a tendency to physical weakness and proneness to emotional maladjustment. Walker (1950) emphasizes the effects of attitudes towards subnormality, considering that parental rejection always exists to some degree even though unexpressed or camouflaged by over-concern and protectiveness. McLachlan (1955) and Evans (1956), however, put more stress on other aspects of the parent-child relationship as sources of difficulties, considering that broken homes, constant discord, illegitimacy, and unrealistic parental ambition

have particularly adverse effects on the E.S.N. child. The effects of failure to meet the subnormal child's needs in school are stressed by Mangus (1950) and Tansley and Gulliford (1960), who consider that the emotional problems of E.S.N. children are often related to the educational and social consequences of low intelligence.

II.—METHODS OF ENQUIRY

In this stage of the enquiry, the most maladjusted children in each of the nine special schools involved (some of the children in the original sample had been transferred to a new special school for junior E.S.N. children) were selected for further intensive study, on the basis of the Bristol Social Adjustment Guide rating and discussion with the headteachers concerned. In all, thirty children in the 'maladjusted' category were studied, and the thirty 'best adjusted' children, matched with the maladjusted pupils in respect of age, sex and school, comprised the control group.

An experienced social worker, appointed specifically to the project, visited the homes of the sixty children to obtain data on the personal and family background of each child. The interviews were conducted in an informal way, which varied according to the circumstances, but the information obtained was recorded in standard and structured form. Whenever possible, information was obtained from both parents, and two or more visits were made to the home when this was considered necessary. All the mothers who had care of their children were seen, as well as 73 per cent. of the fathers, a number of whom stayed at home because they thought that the project might in some way prove helpful to their children. Judgments about parental discipline and attitudes as well as about the personality of the parents were necessarily subjective, though the opinions and experiences of the head teachers, teachers and other professional workers concerned with the families were taken into account in arriving at the final ratings. Apart from the data obtained on each child in the first stage of the investigation, both the writer and the social worker visited the schools to discuss the children with the headteacher.

The sixty children themselves were seen individually at their own schools by the writer on two occasions. Two specifically devised standard oral interview schedules were used, as well as the Bene-Anthony Family Relations Test (1957). The schedules aimed at eliciting the children's main worries, fears and sources of frustration and their attitudes to home, school and self. The questions were framed in simple terms, and mostly required a choice of brief answers in standard form, though a few questions were included which gave scope for freer response.

III.—THE RESEARCH SAMPLE.

As allowance must be made for a certain amount of overlapping between the 'stable,' 'unsettled' and 'maladjusted' categories as defined by arbitrary scores on the Bristol Social Adjustment Guide, the 'most maladjusted' children and the 'best adjusted' children were selected for comparison. In each of these groups of thirty children there were eight junior and twelve senior boys, five junior and five senior girls. The mean score on the Bristol Social Adjustment Guide was 30.20 in the case of the maladjusted group, 6.37 for the control group.

Evidence from other sources confirms this difference in adjustment between the two groups. There was a significant difference in their adjustment in the home situation: seventeen of the maladjusted children, three of the control

group presented problems to their parents (χ^2 14.7, df.1, significant at the .1 per cent. level). Seventeen of the maladjusted group, and again only three of the control group had speech defects, evidence about which was obtained from the interviews with the children and their parents, as well as from the Bristol Social Adjustment Guides; and sixteen of the maladjusted children, six of the controls, had a history of enuresis or soiling (χ^2 7.18, df.1, significant at the 1 per cent. level).

TABLE 1

SOCIOMETRIC STATUS OF CHILDREN IN MALADJUSTED AND CONTROL GROUPS.

Sociometric Status ..	No. of choices received (3×3 criteria)	Number of Children	
		Maladjusted (N=30)	Control (N=30)
Star	15+	4	8
Above average	10-14	—	10
Average	9	2	1
Below average	4-8	14	8
Neglectee	3	2	1
Isolate	0-2	8	1
Not rated (recently joined class) ..	—	—	—

Table 1 shows that there was also a considerable difference in the sociometric status of the children in the two groups. Whereas twenty-four of the maladjusted children were below average in sociometric status (including eight 'isolates' and two 'neglectees'), only ten of the control group were below average, with no more than one 'isolate' and one 'neglectee.'

Eighteen of the maladjusted children showed a predominantly aggressive pattern of symptoms, eight were withdrawn, and in the case of the remaining four there was a mixed pattern of behaviour.

IV.—MAIN FACTORS ASSOCIATED WITH MALADJUSTMENT.

Table 2 shows to what extent various factors were associated with maladjustment in the research sample. While the factors were investigated from the point of view of their aetiological significance, it is realised that a statistically significant association between maladjustment and a particular factor does not necessarily mean that there is a direct causal relationship. It is usually found, too, that a number of factors combine to produce maladjustment, five or more in one case: in twenty-four out of the thirty cases of maladjustment, five or more of the eight groups of factors listed in Table 2 were found in combination.

It will be seen from the table that there were no significant differences between the two groups of children in respect of congenital factors or early developmental difficulties, though in the light of Stott's hypothesis quoted above, it is worthy of note that fourteen of the maladjusted children showed a combination of adverse congenital factors and physical weakness or defect, as compared with eight of the control group. Significant differences between the groups were found in respect of certain physical factors, intellectual and

scholastic factors, and factors connected with the home and especially with the parent-child relationship; these factors are discussed below.

Physical factors.—On the basis of information obtained from the Bristol Social Adjustment Guide, Form 2 H.P., and the mothers of the children, significantly more of the maladjusted children were found to have a record of poor health. Although the difference between the two groups with regard to the incidence of physical defect did not quite reach the 5 per cent level of significance (χ^2 3.78, df.1), in the maladjusted group there were seven children with a history of epilepsy, two spastics, two cases of asthma and one child who had suffered from tuberculosis, whereas none of the children in the control group had had serious defects or illnesses.

All the eight children showing a 'withdrawn' pattern of behaviour disturbance had an unsatisfactory physical history, as compared with two-thirds of the aggressive children. An association between a 'physical deficiency' cluster and over-inhibited behaviour is reported by Hewitt and Jenkins (1946).

TABLE 2
INCIDENCE OF MAIN FACTORS INVESTIGATED.

Factors	No. of children affected		χ^2 (df.1) where significant
	Mal- adjusted Group (N=30)	Control Group (N=30)	
(a) CONGENITAL FACTORS :			
(i) Difficulties during pregnancy	11	9	
(ii) Difficulties at birth	8	7	
(iii) Prematurity (5½lbs. or less at birth) ..	4	8	
History of at least one adverse congenital factor	18	18	
(b) EARLY DEVELOPMENTAL DIFFICULTIES :			
(i) Disturbed feeding in infancy	5	5	
(ii) Hospitalization or institutionalization before 5 years of age	16	16	
History of difficulties in first five years	17	17	
(c) PHYSICAL FACTORS :			
(i) Unsatisfactory physical condition	20	8	9.82*
(ii) Physical defect	13	6	
(iii) Appearance arousing negative reaction	12	2	9.32*
At least one adverse physical factor	25	12	11.9*
(d) INTELLECTUAL AND SCHOLASTIC FACTORS :			
(i) I.Q. below 60	5	3	
(ii) Unsatisfactory progress in reading at special school	14	8	
(iii) Unsatisfactory progress in number at special school	18	5	
(iv) Irregular attendance at school	8	7	11.33*
(v) History of frequent changes of school ..	4	5	
At least one intellectual or scholastic factor	24	16	4.8†

TABLE 2—continued.

Factors	No. of children affected		χ^2 (df.1) where significant
	Mal- adjusted Group (N=30)	Control Group (N=30)	
(e) MATERIAL HOME CONDITIONS AND FAMILY CIRCUMSTANCES :			
(i) Overcrowding	8	3	
(ii) Very low material standards	6	4	
(iii) Father unemployed or in irregular employment	7	7	
(iv) Problem family	2	1	
(v) Financial hardship	5	4	
Adverse material conditions at home	14	12	
(f) PSYCHOLOGICAL FACTORS IN THE HOME :			
(i) Parental instability or unreliability	17	6	8.53*
(ii) Unsatisfactory discipline and attitudes to child	17	6	8.53*
(iii) Lack of appreciation of child's handicap ..	7	1	
(iv) Lack of harmony between parents ..	7	2	
(v) Rejection of child by mother or father ..	4	1	
(vi) Family unhappy in present home	3	3	
(vii) Intellectual standards of family too high for child	2	—	
At least one adverse psychological factor in home	27	12	16.48*
(g) INTERRUPTED OR INCOMPLETE RELATIONSHIPS WITH PARENTS :			
(i) Broken or incomplete home	3	4	
(ii) Child illegitimate	4	—	
(iii) Child adopted	1	1	
(iv) Period in care	4	1	
(v) Substantial period with relatives	3	1	
(vi) Period in residential school or nursery (excluding current residential school) ..	1	1	
(vii) In hospital for 1 month or longer	11	9	
(viii) Frequent periods in hospital	2	—	
Interrupted or incomplete relationships with parents	20	12	4.29†
(h) ILLNESS OF PARENTS :			
I—Mother :			
(i) In poor health	6	3	
(ii) Minor complaints about health	7	6	
(iii) Disabled	1	1	
II—Father :			
(i) In poor health	2	2	
(ii) Minor complaints about health	2	3	
(iii) Disabled	1	—	
Mother unwell or disabled	14	10	
Father unwell or disabled	5	5	

* Significant at 1 per cent. level or beyond.

† Significant at 5 per cent. level or beyond.

As physical appearance was considered likely to affect the child's acceptance by his parents, relatives, peers and teachers, on the grounds that the child who obviously looks subnormal is more apt to arouse feelings of rejection in others, a rating of each child's appearance was made by the writer. The figures given in Table 2 lends some support to this hypothesis.

Intellectual and scholastic factors.—Although the mean I.Qs. of the two groups, as assessed by individual scales, were almost identical (mean I.Q. of maladjusted children 66.93, σ 8.5 ; of control group 67.97, σ 6.4), fewer of the maladjusted children were making satisfactory progress in school. It is, of course, difficult to discover whether unsatisfactory progress in the basic subjects is a causative factor in maladjustment or an effect of maladjustment, but whichever it may be, continued failure to make satisfactory progress in school is likely to cause feelings of frustration in the child and to aggravate any primary disturbance which is present. Overtly, however, the children in both groups expressed little worry, fear or dissatisfaction concerning school, their teachers or school work.

Material and psychological factors in the home.—Psychological factors in the home were much more closely linked with maladjustment in the research sample than were material home conditions. There was, as Table 2 shows, very little difference between the two groups in respect of adverse material conditions at home. Although approximately 23 per cent of the fathers in both groups were in irregular employment and one-half of the families in the total sample were just managing financially or were in difficulties, few of the mothers were in full-time work and the care of the home was generally satisfactory ; over one-half of the sixty families lived in modern homes. The most important factors in the home were the instability or unreliability of the parents and the kind of discipline which they exercised over the child. Inconsistency of discipline and handling was a particularly marked feature in the homes of the maladjusted children. Over-protective and unrealistic attitudes on the part of the parents were also prominent in the sample, but there was relatively little outright rejection of the children.

Table 3, in which the main family relations of the sample, on the basis of the Bene-Anthony Family Relations Test, are shown, demonstrates that the attitudes of the children in the maladjusted group towards their fathers, and the attitudes of the fathers towards the children (as seen by the children themselves) were significantly less positive than in the case of the control group. In both groups, positive attitudes of and towards siblings were outweighed by negative and ambivalent attitudes.

Significantly more of the maladjusted children had had interrupted or incomplete relationships with their parents. There were four illegitimate children in the maladjusted group, none in the control group ; and four of the maladjusted children, one of the control group, had spent a period in the care of the local authority.

There were thirteen eldest children in the maladjusted group, as compared with six in the control group ; there was some evidence that children in this position are given an extra burden of responsibility in the family, particularly as nearly a third of the research sample were members of large families with five or more children.

The attitudes of the parents of the withdrawn children tended to be positive, but discipline was usually weak. In two cases, withdrawal seemed to be the reaction to having siblings or other members of the family who were much brighter. Thirteen of the eighteen aggressive children had a background of insecurity, deprivation, or impatient and intolerant parental attitudes.

TABLE 3

FAMILY RELATIONS OF THE RESEARCH SAMPLE AS SHOWN BY THE BENE-ANTHONY FAMILY RELATIONS TEST

RELATIONS 1251

Attitudes of child towards :	Father			Mother			Siblings		
	Mal. Child- ren	Con- trols	χ^2 df.1	Mal. Child- ren	Con- trols	χ^2 df.1	Mal. Child- ren	Con- trols	χ^2 df.1
N	25	25		27	28		27	28	
Positive	9	17	5.13†	18	23	n.s.	29	35	n.s.
Negative or ambivalent ..	16	8		9	5		53	57	
Attitudes towards child of :	Father			Mother			Siblings		
Positive	5	14	6.88*	17	19	n.s.	28	28	n.s.
Negative or ambivalent ..	20	11		10	9		54	64	

* Significant at 1 per cent. level or beyond.

† Significant at 5 per cent. level or beyond.

Although there had been little difference between the groups in the parents' attitude to their child's transfer to the special school at the time when this was first suggested, about half of each group being hostile towards the idea or very doubtful about it, the parents of the maladjusted children tended to remain fixed in their negative attitude, while some of the parents of the controls became more positive in their attitude towards the special school as the children settled down there.

V.—DISCUSSION.

The above findings, which are supported by a number of other studies which stress the effects of unsatisfactory parental attitudes on the child's emotional development (Kanner, 1953; Beck, 1958; Andry, 1960), underline the need for comprehensive guidance services for the families of educationally subnormal children. The basic financial needs of the families and the physical needs of the children were met in most cases, and the social welfare services were used by a considerable number of the parents. Few of the parents, however, were fully aware of the implications for them of having a subnormal child in the family or had received advice about dealing with the situation. The labelling of their child as 'educationally subnormal' and the subsequent transfer to a special school had been interpreted by most of them as a negative act by 'the authorities.' Only four of the maladjusted children had attended a child guidance clinic.

Although many of the parents will not take the initiative in seeking advice about their child, most of them welcomed the opportunity to discuss the child with the social investigator who called at their homes. It was clear from the interviews that the unsatisfactory attitudes of the parents could be modified to some extent and it seems desirable, therefore, that guidance to

parents should be given by a professional worker visiting the homes of pupils in special schools. This work could, perhaps, be best carried out by social workers appointed to the staffs of special schools, having close links with the school psychological service and other agencies concerned in the welfare of the child.

In addition to guidance services for parents, there is a need for experiments with therapy both within special schools and in a clinical setting. As the interviews showed, E.S.N. children can rarely verbalize their difficulties and in many cases are not fully aware of them. Direct interview-therapy with the children, therefore, is unlikely to be successful in the majority of cases, though experimental individual and group therapy of this kind with the older and less dull pupils would be of interest. Activity group therapy (Slavson, 1952), modified to suit children of limited intelligence, seems a more suitable form of treatment for many maladjusted E.S.N. children, as Evans (1956) demonstrated. Mundy (1957) suggests that even severely subnormal children who are not psychotic or organically impaired can be helped by psychotherapy.

Smaller classes in special schools for the E.S.N. would help teachers both to cope with behaviour problems and to give more individual attention to pupils. The close link between maladjustment and low attainment shows that a number of children continue, even at the special school, to fail to derive satisfaction from a sense of progress and achievement. The 1964 Report of the Department of Education and Science records that in January, 1963, 13.6 per cent. of the pupils in special schools for the E.S.N. were in 'oversize' classes, that is, in classes of over twenty children, and three-quarters of the classes contained sixteen or more children. In schools which have a high proportion of maladjusted children, twelve children in a class would be a more desirable maximum number, especially in the light of the Underwood Report (Ministry of Education, 1955) which, discussing the educational needs of maladjusted children, concluded that ordinarily a teacher cannot satisfactorily meet the needs of more than ten maladjusted children.

The prevention of situations conducive to maladjustment might be helped by the establishment of special centres as in the United States (Giannini, 1957; Kanner, 1961), where parents of subnormal children can obtain expert advice at an early stage. Such centres might also encourage systematic research into the many problems connected with educationally subnormal children which are still awaiting attention.

ACKNOWLEDGEMENTS.—The writer is particularly indebted to Miss Elisabeth Gruber, Research Officer to the project, who was responsible for carrying out the home visits and who contributed much to the analysis and interpretation of the data. He would also like to thank the following for their help and co-operation: the Directors of Education and Medical Officers for Cardiff, Carmarthenshire, Cardiganshire, Monmouthshire, Newport and Swansea; the Children's Officer for Cardiff; the head teachers and staff of the special schools concerned; Dr. D. H. Stott, for helpful comments on this paper; Miss Margaret Lewis, secretary to the project; and the Department of Education and Science, for their interest and financial support.

VI.—REFERENCES.

- ANDRY, R. E. (1960). *Delinquency and Parental Pathology*. London: Methuen.
 BECK, MARIGOLD A. (1958). *Some Environmental Factors in Mental Retardation*. Ph.D. Thesis, University of Cambridge.

- BENE, EVA, and ANTHONY, J. (1957). *Family Relations Test*. N.F.E.R.
- CHAZAN, M. (1964). The incidence and nature of maladjustment among children in schools for the educationally subnormal. *Brit. J. Educ. Psychol.*, 34, 292-304.
- DEPARTMENT OF EDUCATION AND SCIENCE (1964). *Statistics of Education*, Part I, 1963. London: H.M.S.O.
- EVANS, D. (1956). *An Experimental Study of a Group of Seriously-Maladjusted Educationally Subnormal Children*. M.A. (Education) Thesis, University of Birmingham.
- GIANNINI, MARGARET J. (1957). Diagnostic approach in mental retardation. *Public Health News*, 38, 10, 322-332.
- HEWITT, L. E., and JENKINS, R. L. (1946). *Fundamental Patterns of Maladjustment: the Dynamics of their Origin*. Illinois: Green.
- KANNER, L. F. (1953). Parents' feelings about retarded children. *Amer. J. Ment. Defic.*, 57, 375-383.
- KANNER, L. F. (1961). Parent counseling: in *Mental Retardation*, ed. Rothstein, J. H., New York: Holt, Rinehart and Winston.
- McLACHLAN, D. G. (1955). Emotional aspects of the backward child. *Amer. J. Ment. Defic.*, 60, 323-330.
- MANGUS, A. R. (1950). Effect of mental and educational retardation on personality development of children. *Amer. J. Ment. Defic.*, 55, 208-212.
- MINISTRY OF EDUCATION (1955). *Report of the Committee on Maladjusted Children*. London: H.M.S.O.
- MUNDY, LYDIA (1957). Therapy with physically and mentally handicapped children in a mental deficiency hospital. *J. Clin. Psychol.*, 13, 3-9.
- SLAVSON, S. R. (1952). *Child Psychotherapy*. New York: Columbia University Press.
- STOTT, D. H. (1957). Physical and mental handicaps following a disturbed pregnancy. *Lancet*, i: 1006-12.
- STOTT, D. H. (1959). Evidence for pre-natal impairment of temperament in mentally retarded children. *Vita Humana*, 2, 125-148.
- TANSLEY, A. E., and GULLIFORD, R. (1960). *The Education of Slow Learning Children*. London: Routledge and Kegan Paul.
- WALKER, G. H. (1950). Social and emotional problems of the mentally retarded child. *Amer. J. Ment. Defic.*, 55, 132-138.

(Manuscript received 12th April, 1965.)

ACHIEVEMENT MOTIVATION AND GOAL-SETTING BEHAVIOUR IN THE CLASSROOM*

By BISHWA NATH MUKHERJEE

(Indiana University)

SUMMARY. The study demonstrates the construct validity of a forced-choice test called *Sentence Completion Test* as a measure of achievement motivation. Four hypotheses derived from McClelland's studies on achievement motivation and level of aspiration were tested. Excepting one, all the hypotheses were confirmed. It was found that Ss with high scores on SCT showed more positive goal discrepancy scores than others. A multivariate test of analysis of variance also revealed a significant over-all difference between the high and low *n* Ach groups in terms of three different measures derived from the estimated, aspired and actual scores on three psychology quizzes given to eighty-five college freshmen.

I.—INTRODUCTION.

IN recent years, the relationship between achievement motivation (*n* Ach) and goal setting behaviour has been studied for a number of non-learning tasks by McClelland and his associates (McClelland *et al.*, 1953, p. 242-248; Atkinson and Reitman, 1956; Atkinson, 1950; Atkinson, 1958, p. 288-305; Clark, Teevan and Ricciuti, 1956). More recently, the relationship between them has also been studied in a learning situation (Kausler and Trapp, 1958; Ricciuti and Schultz, 1958). Most of these studies show that people with strong generalized achievement motivation usually set higher goals than those with low *n* achievement, particularly when the reality determinant in the level of aspiration situation is unstructured. In the present study, the same hypothesis has been tested in order to establish the construct validity of a forced-choice questionnaire, called *Sentence Completion Test* (SCT). The SCT has been recently proposed as a reliable and valid tool for the measurement of achievement motivation (Mukherjee, 1965a).

The study primarily involved a comparison between Ss with high scores on SCT and Ss with low scores on SCT, with regard to the level of performance which they explicitly undertook to reach in three psychology examinations. It was the purpose of the study to test the null hypothesis that there is no difference between the goal setting behaviour of the high *n* Ach group and low *n* Ach group. More specifically, the experimental hypotheses investigated in this study were: (a) Ss with high scores on SCT will show more positive "goal discrepancy scores" (Eysenck and Himmelweit, 1946) than Ss with low scores on SCT, (b) the difference between the two groups in terms of goal discrepancy scores will be more clear cut on the first trial (session) of goal setting than the subsequent ones when Ss are aware of their progress, (c) Ss scoring high on SCT will tend to raise their aspiration level when their actual performance is relatively poor. Those performing well, on the other hand, will show a positive but a relatively lower goal discrepancy score compared to the poor performers. Ss scoring low on the SCT, on the other hand will show no such shift in their goal setting irrespective of their actual performance on the psychology examination. As such, there will be an inverse relationship between

* The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare under the provisions of the Co-operative Research Program.

goal discrepancy score and actual performance and this will be found only in the high *n* Ach group and *not* in the low *n* Ach group, and (d) in terms of "Judgment Discrepancy Scores" (Eysenck, 1952), Ss with high SCT scores will score lower than Ss with low score on SCT.

The last hypothesis is borne out of the observation that people with high achievement motivation tend to rate themselves low in some achievement related traits, such as intelligence (Martire, 1956; Ricciuti and Schultz, 1958). They seem to be dissatisfied with their performance and therefore express their judgment regarding their achievement in a more restrained manner. If this is true, then it seems conceivable that Ss scoring high on SCT will have high level of aspiration (high positive goal discrepancy score) and will tend to underestimate their past performance (high negative judgment discrepancy score). Thus, the correlation between goal and judgment discrepancy scores is expected to be negative in the case of high *n* Ach group. No such predictive statement can be made regarding the low *n* Ach group.

In addition to the goal discrepancy and judgment discrepancy scores, the present study also employed the Zelen (1955) rigidity index, which consisted of an additive combination of the total number of shifts during the course of the experimental period with the absolute total amount of shifts on all the trials. The higher the score on this index, the more flexible the individual is in his goal setting behavior. This flexibility measure was preferred to Eysenck's (1947) Index of Flexibility because the former takes into consideration not only the amount an individual shifts his level of aspiration as in the case of Eysenck's index but also the number of shifts in the positive as well as negative directions. The measure was solely utilized to investigate if *n* Ach has any relationship with flexibility in a level of aspiration situation.

II.—METHOD.

Subjects.

The Ss were forty-three male and forty-two female students enrolled in two sections of an introductory psychology course at the South-Eastern Campus of Indiana University, Jeffersonville, Indiana, during the fall 1962 semester. Almost 95 per cent. of them were freshmen. Their ages ranged between 18 and 22 years.

Instrument.

The *Sentence Completion Test* developed by Mukherjee (1965a) was used to classify Ss in two criterion groups. The test-retest reliability of the test after an interval of nearly two months was found to be .71 for another sample of eighty-seven freshmen. The coefficient was almost .83 for a sample of fifty-eight students more heterogeneous in population characteristics than the sample of the present study. The Kuder-Richardson reliability estimate was .716 for a sample of 248 college students. The correlation of SCT scores and scores on the *n* Ach scale composed of seventeen items borrowed from Murray (1938, p. 165) turned out to be .44 for a sample of fifty-eight students. The correlation is significant at the .01 level. Evidence of concurrent as well as construct validity of the test is borne out of a few more correlational studies, a laboratory investigation which successfully predicted better performance on a test of perceptual speed for Ss scoring high on SCT than those scoring low and finally a study on the relationship between achievement motivation as measured by SCT and several dimensions of self concept (Mukherjee, 1966).

The Ss of the present study took the SCT during the regular class periods two weeks before the level of aspiration study began. They were told that the purpose of testing was to obtain certain information regarding the uniqueness of their behaviour which could be used for the purpose of counselling. The forty-three Ss scoring seventeen or above on the SCT were arbitrarily designated as the high n Ach group; and the remaining forty-two Ss having scores less than seventeen formed the low n Ach group. In the high n Ach group, there were twenty females. The number of females in the low n Ach group was twenty-two only.

The same Ss were also given the Vocabulary Test of Thurstone's (1938) PMA battery at the very beginning of the semester. The Vocabulary Test has been found to predict satisfactorily the total grade points in an Introductory Psychology course for a comparable sample (Mukherjee, 1965*b*). It was used in the present study mainly to see if there is any intrinsic correlation between level of aspiration scores and need achievement scores beyond what is induced by their common dependence upon verbal intelligence.

Procedure.

The entire study was conducted in the regular class hours. It continued over almost two and half months. During this period, all Ss had three psychology quizzes as part of their regular evaluation. The total score in each of these quizzes was 100. All the quizzes consisted of multiple choice and matching exercises. Immediately before the scores of each S were announced in the classroom, he was required to fill out a small questionnaire which among other things included the following items:

- (1) What score had you in the last psychology quiz?
- (2) How satisfied are you with your performance in the last quiz?
- (3) Estimate what would be your score this time.

After the Ss filled out the questionnaire and score of each student was announced publicly in the classroom, they were given the following instructions: "On the next quiz that you are going to take after three weeks, the maximum score will be 100 as usual. Please state in writing at the bottom of the questionnaire what score you *expect* to make on the next examination. Your answer should be in terms of percentage and *not* letter grade."

The same procedure was followed in the second and third sessions of the experiment. Before each feed back, the Ss estimated the scores they were going to receive and after the feed back, they indicated their next level of performance on the psychology examination.

From the estimated and aspired scores, goal discrepancy, judgment discrepancy, and flexibility scores were derived for Ss of high and Ss of low n Ach groups. These scores were then subjected to a multivariate analysis of variance for testing the equality of mean vectors*. In one analysis all the seven measures were used. In another, only three goal discrepancy measures, one obtained for each level of aspiration situation, were used. Product moment intercorrelations among all the measures were computed separately for both the high and low n Ach groups. In order to determine further to what extent the Ss of these two groups can be differentiated in terms of their profiles provided by seven level of aspiration scores, Fisher's (1936) discriminant function analysis was applied to the obtained data. This was also done separately for the three-goal discrepancy scores for testing our second hypothesis.

* Most of the computations for this study were performed on an IBM 709 Computer located at Indiana University, Bloomington. The facilities obtained in this connection from the Research Computing Center of Indiana University are gratefully acknowledged. Special appreciation is due to Miss Ruth Kobbe for help in the data processing.

III.—RESULTS.

A comparison of the means and variances for the two groups is provided by means of Table 1. The intercorrelations among all the measures used in the present day study are reported separately for high and low *n* Ach groups in Table 2. It may be noted that the high intercorrelations among the psychology quiz scores reflect a satisfactory amount of reliability of the psychology examinations.

TABLE 1

GROUP MEANS AND VARIANCES FOR HIGH AND LOW *n* ACH. GROUPS, THE DIFFERENCES BETWEEN THE MEANS AND VALUES OF *t* FOR DIFFERENCES BETWEEN GROUP MEANS ON ALL MEASURES.

Variable	Group Means		Difference	Group Variances		<i>t</i> *	Prob.**
	High <i>n</i> Ach (N=43)	Low <i>n</i> Ach (N=42)		High <i>n</i>	Low <i>n</i>		
Goal Dis. I	16.023	2.095	13.928	135.261	69.893	6.351	.0002
Goal Dis. II	15.256	8.833	6.423	149.385	243.215	2.110	.0250
Goal Dis. III	19.326	10.333	9.993	175.415	147.203	3.630	.0004
Judge Dis. I	12.535	1.809	10.726	95.588	68.060	5.474	.0003
Judge Dis. II	1.698	1.000	0.698	176.073	197.415	0.235	—
Judge Dis. III	0.163	2.071	-1.908	144.425	118.556	-0.768	—
Flexibility	17.628	25.809	-8.181	164.143	337.670	-2.374	.0100

*—Calculated under the assumption of unequal variances.

**—One-tailed test.

From the intercorrelations obtained for the entire sample, a few first-order partial correlations were computed to eliminate the effect of verbal intelligence. The partial correlations between SCT scores and Mean Goal Discrepancy was found to be .41; between SCT and Judgmental Discrepancy, .42 and between SCT and Flexibility, -.28, all significant at the .02 level.

The multivariate test statistic computed for the equality of mean vectors for the high and low *n* Ach groups was Hotelling's (1931) T^2 . The obtained T^2 was 58.809. In order to test its significance, the T^2 was converted into F statistic. The obtained result clearly showed that the overall difference between the two criterion groups in terms of all the seven level of aspiration measures was statistically significant beyond .001 level ($F=7.809$ with 7 and 77 d.f.).

The discriminant function coefficients obtained for both the analyses appear in Table 3. In Table 4, the means and variances of the discriminant scores for both the groups are displayed separately for the two analyses together with the corresponding Mahalanobis (1936) D^2 statistics. It can be seen from Table 3 that the largest contribution to the discriminant function obtained with all the seven variables is provided by the first judgment discrepancy score and then the first goal discrepancy score both with positive weights. When only three goal discrepancy scores were used in the analysis, it turned out that the difference between the high and low *n* Ach groups in terms of Mahalanobis D^2 was significant beyond .001 level as can be seen from Table 4. The discriminant function for this analysis reported in Table 3 also showed that in separating the two *n* Ach groups, it was the first goal discrepancy score that played the major role which directly supports our second hypothesis.

TABLE 2
PRODUCT MOMENT INTERCORRELATIONS* AMONG LEVEL OF ASPIRATION MEASURES AND
ACTUAL SCORES ON PSYCHOLOGY QUIZES SHOWN SEPARATELY FOR HIGH AND LOW "n"
ACHIEVEMENT GROUPS.

Variable	GD I 1	GD II 2	GD III 3	JD I 4	JD II 5	JD III 6	Flexib 7	Psychology Quiz Score				Mean GD 12	Mean LD 13	Total Psy. 14
								First 8	Second 9	Third 10	Vocab Test 11			
Goal Discrep. I	1.000	.378	.328	.776	-.053	-.040	-.035	-.519	-.310	-.341	.098	.750	.307	-.416
Goal Discrep. II318	1.000	.265	.220	.450	.182	.212	-.322	-.486	-.244	.103	.733	.494	-.371
Goal Discrep. III474	.404	1.000	.121	.268	.222	.355	-.395	-.288	-.577	-.032	.739	.355	-.456
Judge Discrep. I550	.367	.322	1.000	-.212	.009	.040	-.481	-.241	-.231	-.011	.485	.339	-.336
Judge Discrep. II	-.115	.351	.073	.120	1.000	.227	.009	-.100	-.346	-.232	-.007	.307	.667	-.243
Judge Discrep. III034	.185	.181	-.117	.349	1.000	.456	-.194	-.286	-.318	-.035	.171	.724	-.287
Flexibility089	.299	.191	.349	.275	.111	1.000	-.256	-.242	-.373	-.097	.251	.287	-.314
Psych. Quiz I	-.165	-.196	-.226	-.269	-.159	.127	.176	1.000	.825	.809	.418	-.554	-.401	.937
Psych. Quiz II024	-.253	-.191	-.067	-.324	-.275	-.009	.711	1.000	.791	.521	-.487	-.497	.930
Psych. Quiz III	-.052	-.181	-.306	-.134	-.262	-.265	-.045	.745	.849	1.000	.472	-.531	-.439	.933
Vocabulary Test	-.038	-.073	.019	.007	-.013	.101	.081	.391	.529	.483	1.000	-.051	.047	.531
Mean Goal Discr.680	.822	.798	.505	.191	.191	.496	-.255	-.215	-.247	.093	1.000	.521	-.561
Mean Judge Discr.156	.475	.269	.252	.802	.705	.378	-.277	-.385	-.365	.019	.425	1.000	-.477
Total Psych. Score	-.065	-.228	-.264	-.163	-.274	-.247	.077	-.877	.932	.947	.512	-.259	-.375	1.000

* Correlations for the High "n" Ach group appear above the main diagonal; for the low group, they are given below the main diagonal. $r = .38$ required for significance at .01 level.

TABLE 3

DISCRIMINANT FUNCTION COEFFICIENTS FOR EACH OF THE VARIABLES INCLUDED IN TWO ANALYSES.

Variable	Discriminant Coefficients	
	Analyses involving all Variables	Goal Discrepancy
Goal Discrepancy I000634	.001549
Goal Discrepancy II000293	— .0046
Goal Discrepancy III000424	.000210
Judgment Discrepancy Score I001146	—
Judgement Discrepancy Score II000191	—
Judgement Discrepancy Score III	— .000080	—
Flexibility Score000772	—

TABLE 4

MEAN AND VARIANCE FOR THE TWO GROUPS IN TERMS OF DISCRIMINANT FUNCTION SCORES TOGETHER WITH THE DISTANCE BETWEEN THEM AS FOUND IN TWO SEPARATE ANALYSES.

Group	Discriminant Function Analysis Involving			
	All Seven Measures		Goal Discrep. only	
	Mean	Variance	Mean	Variance
High <i>n</i> Ach Group02388	.0004955	.02818	.000357
Low <i>n</i> Ach Group	— .00953	.0003073	.00501	.000199
Mahalanobis D ² Statistic	2.773		1.923	
F Statistic	7.809**		13.293**	
Degrees of Freedom	(7,77)		(3,77)	

** $p < .01$.

IV.—DISCUSSION.

The significant difference in the predicted direction between the high and low scorers on the SCT clearly demonstrates the construct validity of the proposed forced-choice test as a measure of achievement motivation. The marked tendency of the high group to express higher goal setting behaviour as measured by the goal discrepancy scores, particularly in the first session, also suggests that the relationship between achievement motivation and S's expressed expectancy of performance is clearly positive when the reality determiners of aspiration level operate in minimum. This finding is totally in accord with that of Kausler and Trapp (1958) who found that the difference between high and low *n* Ach groups in goal discrepancy scores tends to diminish with practice on a digit symbol task when Ss were informed of their progress.

The intercorrelation between the mean goal discrepancy scores and the mean psychology quiz scores is $-.56$ for the high n Ach group. The corresponding value for the low n Ach group is only $-.26$ which is not significant even at $.05$ level (Two-tailed test). We thus find that our third hypothesis is also confirmed.

In examining the relationship between the mean judgment discrepancy scores and the mean goal discrepancy scores, we find that the correlation between the two scores are almost the same for both high and low n Ach groups. This finding in conjunction with the mean judgment discrepancy scores for the successive sessions for both groups totally discard the fourth hypothesis we advanced. However, the results do indicate that with successive feed backs, those in the high n Ach group began underestimating their past performance whereas those in low n Ach group showed a tendency to overestimate. Further studies extending over a longer period of time seem necessary to substantiate this interesting observation.

The higher flexibility score on the part of Ss scoring low in SCT seems to be another interesting finding of this study. The product moment correlation between the flexibility scores and SCT scores was found to be $-.296$ which indicates a negative relationship between n Ach and flexibility. This finding is in accord with the result obtained by Mukherjee (1966) who found a low but negative correlation ($r = .136$) between SCT scores and self ratings on flexibility dimension. The result is interpreted to mean that people with high n Ach tend to show more determination, perseverance in the pursuit of their high goals and as such they are less flexible in terms of their shifts in stated goals. The results of this study clearly indicate some interesting relationships between anticipation of achievement and differential achievement in college. In addition the study demonstrates the usefulness of applying multivariate statistical techniques to the analysis of goal setting behaviour. The discriminant function analysis was used in the present study not for the purpose of classifying Ss in terms of high or low n Ach but it provided a useful means of specifying the nature of differences between the two groups in terms of the various aspects of goal setting behavior that can be studied longitudinally.

Methodologically, the study supports the usefulness of the *Sentence Completion Test* for differentiating Ss on the n Ach variable. Since the administration time of this group forced-choice test is approximately 25 min. and its scoring is quick and objective and above all very little faking is possible in this test, it seems quite promising as a research tool.

V.—SUMMARY AND CONCLUSIONS.

The primary purpose of this research was to investigate the relationship between several level of aspiration measures derived from the estimated, aspired, and actual scores on three successive psychology quizzes and achievement motivation scores obtained by means of a forced-choice test called *Sentence Completion Test*. Eighty-five college freshmen who took the SCT were split into two groups on the basis of their scores on SCT. It was hypothesized that Ss scoring high on the SCT will show higher positive goal discrepancy scores compared to Ss of the low group. The hypothesis is confirmed from the mean scores of the two groups on three level of aspiration situations. It was also found that the differences between the two groups in terms of goal discrepancy score as well as judgment discrepancy score were more clear cut on the first goal setting session. Other interesting findings of this study are discussed in the light of the construct of n Ach. The study demonstrates the construct validity of SCT as a measure of need achievement.

VI.—REFERENCES.

- ATKINSON, J. W. (1950). Studies in projective measurement of achievement motivation. Unpublished doctoral dissertation, University of Michigan.
- ATKINSON, J. W., and REITMAN, W. R. (1956). Performance as a function of motive strength expectancy of goal attainment. *J. Abnorm. Soc. Psychol.*, 53, 361-366.
- ATKINSON, J. W. (1958). *Motives in Fantasy, Action and Society*. New York: Van Nostrand and Co.
- CLARK, R. A., TEEVAN, R., and RICCIUTI, H. N. (1956). Hope of success and fear of failure as aspects of need for achievement. *J. Abnorm. Soc. Psychol.*, 53, 182-186.
- EYSENCK, H. J., and HIMMELWEIT, H. T. (1946). An experimental study of the reactions of neurotics to experiences of success and failures. *J. Gener. Psychol.*, 35, 59-75.
- EYSENCK, H. J. (1947). *Dimensions of Personality*. London: Routledge and Kegan Paul.
- EYSENCK, H. J. (1952). *The Scientific Study of Personality*. London: Routledge and Kegan Paul.
- FISHER, R. A. (1936). The use of multiple measurements in taxonomic problems. *Ann. Eugen.*, 7, 179-188.
- HOTELLING, H. H. (1931). The generalization of Student's ratio. *Ann. Math. Statist.*, 2, 360-378.
- KAUSLER, D. H., and TRAPP, E. P. (1958). Achievement motivation and goal-getting behavior on a learning task. *J. Exp. Psychol.*, 55, 575-578.
- MAHALANOBIS, P. C. (1936). On the generalized distance in Statistics. *Proc. Nat. Inst. Sc. (India)*, 12, 49-55.
- MARTIRE, J. G. (1956). Relationships between the self concept and differences in the strength and generality of achievement motivation. *J. Pers.*, 19, 395-399.
- MCCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., and LOWELL, E. L. (1953). *The Achievement Motive*. New York: Appleton-Century-Crofts.
- MUKHERJEE, B. N. (1965a). A forced-choice test of achievement motivation. *J. Indian Acad. Appl. Psychol. Sept.*
- MUKHERJEE, B. N. (1965b). The prediction of grades in introductory psychology from tests of primary mental abilities. *Educ. Psychol. Measmt.*, 25, 557-564.
- MUKHERJEE, B. N. (1966). Self perception and achievement motivation. *J. Psychol. Res.* (in press).
- MURRAY, H. A. (1938). *Explorations in Personality*. New York: Oxford University Press.
- RICCIUTI, H. N. and SCHULTZ, D. G. (1958). Level of aspiration measures and self estimates of personality in relation to achievement motivation. *Technical Report, Office of Naval Research Contract Nonr.—694 (00)*. Princeton, N.J.: Educational Testing Service (January).
- THURSTONE, L. L. (1958). Primary mental abilities. *Psychometric Monogr.*, No. 1..
- ZELEN, S. L. (1955). Goal-setting rigidity in an ambiguous situation. *J. Consult. Psychol.*, 19, 395-399.

(Manuscript received 30th November, 1964.)

A COMPARATIVE STUDY OF THE ATTITUDES OF WEST AFRICAN STUDENTS AND TEACHERS TOWARDS THE MODERN APPROACH IN TEACHING

By OSCAR FERRON

(*Ahmadu Bello University, Zaria, N. Nigeria*)

SUMMARY. Two experiments were carried out in order to determine the facility with which attitudes of West African students and teachers towards the 'modern' approach in teaching can be expected to change. The first experiment involved teachers in Sierra Leone and Nigeria, who had been trained at the University College of Sierra Leone. Their attitudes towards the modern approach in the classroom were compared with those of groups of British, American and African teachers in Sierra Leone and England by means of a Likert type attitude scale. A statistically significant favourable difference was found between their responses and those of all other groups with whom they were compared. In the second experiment, graduate and under-graduate groups of students in Zaria responded to the same scale at the commencement and end of their respective courses in Education. All groups showed a significant gain over their previous scores.

I.—INTRODUCTION.

In this investigation we were primarily interested in teachers' private opinions, whether or not they had the courage of their convictions. When the ground has been prepared by the adoption of the kind of private opinions regarding modern classroom procedures which we envisage, the force of public opinion will inevitably lead to a certain set of desirable actions. Of course, there are bound to be inconsistencies, but it is enough for our purpose that teachers hold certain views which they are prepared to divulge to others. Moreover, if changes in classroom procedures in West Africa are desired, it is obviously a necessary first step to discover how African teachers feel about these changes, and if their attitudes are unfavourable, to devise ways and means whereby they can be influenced to become more favourably disposed towards them.

II.—METHOD.

It was not possible to proceed according to the usual experimental design, using experimental and control groups, because of the numbers of subjects involved. At Freetown there were never more than about a dozen students each year, who were enrolled for the post-graduate diploma course, hence it was decided to test all former students who were available, and who had previously attended the course over a period of three consecutive academic sessions.

The next problem was to devise a suitable scale. The two most widely used scales for the measurement of attitudes are the 'Likert' and 'Thurstone' scales, respectively. The latter is probably the most objective method yet devised, but it is not always convenient to use, since its initial construction depends upon the services of a large number of judges. On the other hand the 'Likert' method has much to recommend it. Not only is it easy to construct, but if the items are judged to be internally consistent on the basis of the investigator's experience, it has proved itself an effective measuring instrument.

Thus according to Newcomb: "It has been shown that . . . the split half reliability coefficients are higher when scored by Likert than by Thurstone methods." Likert himself has shown that his simpler scales measuring attitudes relating to international relations, when compared with the Thurstone-Droba war scale, "yielded the same reliability coefficients with practically one half as many items." He introduced a method of weighting the graded responses on any item according to their centroid or z-values, but it was discovered later that the simpler technique of scoring each item 5, 4, 3, 2, 1 gave practically the same results, and so the more complicated procedure was dropped. In the present investigation each item was scored +2, +1, 0 according to the progressive, neutral or traditional nature of the response, no distinction being made between 'Strongly agree' and 'Agree,' and between 'Strongly disagree' and 'Disagree.'

Since the scale was intended to discriminate between teachers with a progressive and a traditional outlook respectively, 100 items were initially devised, agreement in the case of forty suggesting progressiveness, and agreement on the remaining sixty denoting a traditional outlook. Normally there should be a 50:50 balance, but it was felt that people with an authoritarian or traditional outlook usually like it to be known that they are 'progressive,' and consequently there might be some concealment of true attitude. On the other hand, some people who feel rather strongly against the so-called 'new fangled methods' could be expected to check traditional items quite sincerely. Items were drawn up on the basis of the following dichotomy:

TABLE 1
BASIS OF THE LIKERT SCALE.

No.	Traditional Theories and Techniques.	Modern Theories and Techniques.
1.	Doctrine of Formal Discipline.	Doctrine of Interest.
2.	Lecture method-telling the children everything.	Posing problems-drawing the children out by question and answer.
3.	Techniques in which examination considerations are paramount.	Techniques in which the lesson is related to the children's own experiences and interests.
4.	Techniques involving verbalism and abstract concepts.	Techniques involving the use of audio-visual aids—proceeding from the concrete to the abstract.
5.	Teaching <i>en masse</i>	Individual methods.
6.	Use of corporal punishment.	No corporal punishment in the teaching situation.
7.	Education as a means of imparting information.	Education as a training for life.

The questionnaire was sent to all secondary school teachers in Sierra Leone under a covering letter addressed to each by name. Friends and former colleagues in England were also approached. Eventually 216 questionnaires representing various categories of Sierra Leonean, British and American teachers, were completed. The 'Upper and Lower Thirds' technique was

used to calculate discriminating values for each item. The mean maximum score possible for each item was 2, the minimum 0. The difference 2-0 gave the maximum index possible. To make a selection of the best items, one could have drawn the line at a discriminating value of .8. This would have yielded twenty-nine items. However, it was felt that it would be a definite advantage if one could retain the dichotomy outlined above. Hence forty of the best items spread over these categories were eventually selected. The revised scale is shown as an appendix to this paper.

The scripts were re-assessed on the basis of the purified scale, and the split-half reliability coefficient calculated. When corrected by the Spearman-Brown formula, value of .86 ($N=216$) was obtained. The split-half reliability coefficient often tends in attitude research to be lower than the test re-test coefficient, and in any case, as Vernon points out: "It is probably a mistake to aim at extremely high reliability in attitude scales, since this may be obtained at the expense of validity."

However, the problem of validity in psychological testing remains unresolved. A test is said to be valid to the extent that it measures what it is supposed to measure. Attitudes are predispositions which we become aware of through behaviour, but when the nature of attitude is not yet precisely understood, and the behaviour to which it gives rise may be distorted for reasons of courtesy or expediency, it makes it very difficult indeed to guarantee the validity of any particular attitude scale. Moreover, the assessment of validity is in the last resort, a subjective one. There has never been a psychological test that has yet been devised, the validity of which can be guaranteed independently of the impressions and reactions of a human personality. It seems feasible, therefore, to measure behaviour as an index of attitude, and then define the attitude as 'what the test measures.' The validity of a test must inhere in its construction, and must inevitably be based on a subjective assessment of its various items in relation to a logical analysis of the trait in question, as it exists in the investigator's perception. Thus McNemar has pointed out that: "A scale for attitude toward an issue is said to be valid because the items in the scale make it so, that is the items have been judged to be valid." On this assumption we are reasonably satisfied that we have produced a valid measure. As a very rough check, the author carefully considered the responses from former colleagues whom he had known from two to ten years or more. He found no discrepancies: the progressive ones obtained high scores, while those with a more traditional outlook obtained lower scores. On the other hand, as we are not concerned with an immediate action orientation, obviously this question of validity is not of overriding importance. Remmers has in fact suggested that: "If we are interested only in knowing what the present attitudes of a given group are, we can equate validity with reliability."

III.—RESULTS.

The scripts were divided up into lots representing the following categories of teachers, whose mean scores are given in the table below. Numbers in brackets indicate the number of teachers responding:

(a) Former Dip. Ed. Students (twenty-one).....	67.4
(b) American Graduates in Sierra Leone (twenty-nine)	55.8
(c) British Graduates in Sierra Leone (twenty-six).....	53.0
(d) British Teachers in England (fifty-one)	51.0
(e) African Teachers (Graduates, U.K. Teachers Cert., Teachers Advanced Certificate, S.L.) (twenty-nine)	49.0

To summarise the findings without going into details, the following points may be mentioned :

- (1) The differences in the means between the former Dip.Ed. students and all other groups were statistically significant.
- (2) The difference between African and American graduates in Sierra Leone was found to be significant at the $P = .2$ level only.
- (3) No significant difference was found between British teachers in England and African teachers of a comparable standard of education in Sierra Leone.
- (4) There was no significant difference between British teachers in England and British teachers in Sierra Leone.
- (5) There was no significant difference between British graduates and African graduates in Sierra Leone.

In October, 1963, the scale was administered to the following groups of students at the commencement of their respective courses in Education in Zaria:

- | | |
|--|-------------|
| (a) Post-Graduate Art Teachers Certificate | 10 students |
| (b) Physical Education Diploma, 2nd Year | 11 " |
| (c) B.A./B.Sc. (Ed.), 1st and 2nd Years | 33 " |
| (d) Nigerian Certificate in Education | 119 " |

Towards the end of their courses the questionnaire was administered again. There was a significant gain for each group. The level of significance is shown in the table below :

TABLE 2
STATISTICAL NATURE OF GAINS-ZARIA GROUPS.

Group	Level of Significance $P = .001$
A.T.C.	$P = .05$
P.E.D.	$P = .001$
B.A./B.Sc.	$P = .05$
N.C.E.	

In spite of the above results, however, it was felt that in the main, and especially in the case of the small intimate post-graduate groups in Freetown and Zaria, the changes may have been due to :

- (a) the students' orientation towards giving the right answer, and/or
- (b) a desire to please.

In order to investigate this matter, a discussion was initiated as a result of a student's question during a lecture. It lasted an hour, and was as natural and spontaneous as any in the author's experience. The students did almost all the talking, and although their powers of expression were weak, the maturity of their ideas, and their convictions were quite impressive. The investigator made notes during the discussion, and while it was not possible to reproduce the exact words uttered, an attempt was made to recapture the gist of what was said by writing the reports immediately after the students had dispersed. These reports are sufficient evidence in themselves to indicate that a genuine change had occurred in the case of the students who spoke.

IV.—CONCLUSION.

It is possible to argue theoretically that as West African students and teachers have inherited a somewhat rigid and authoritarian way of life, it would be very difficult to bring about the desired changes in their attitudes towards the 'modern' classroom approach. But in the face of the evidence presented above, it is equally difficult not to be convinced that such changes among some of the respondents, at any rate, are genuine. However, it is impossible at this stage to say precisely to what these changes were due, since several lecturers and tutors with varying approaches, were involved. One thing appears to be certain, and that is that we are not as seriously restricted by present cultural conditions in West Africa as has been suggested. One possible explanation is that, despite his centuries old authoritarian way of life, the educated African has a definite orientation towards becoming more and more civilized, which, in effect implies the assimilation of Western cultural values and procedures. But the solution to the whole problem of assimilation would seem to lie in adequate communication, and a satisfactory emotional interaction between those responsible for the training of teachers and their students.

V.—REFERENCES.

- FERRON, O. (1957). *The Fostering of International Understanding with Special Reference to the Science Course in a Secondary Technical School*. M.A. Thesis, London University.
- FERRON, O. (1964). *A Study of Certain Factors Affecting the Tested Intelligence of Some Groups of West African Children*. Ph.D. Thesis, London University.
- LIKERT, R. (1932). Technique for the measurement of attitudes. *Arch. Psychol.*, Vol. 22, No. 140.
- McNEMAR, Q. (1946). Opinion-attitude methodology. *Psychol. Bull.*, 43, 289, 374.
- NEWCOMBE, T. M. (1952). *Social Psychology*. London: Tavistock.
- REMMERS, H. H. (1954). *Introduction to Opinion and Attitude Measurement*. New York: Harper.
- VERNON, P. E. (1957). *Personality Tests and Assessments*. Methuen.

(Manuscript received 27th October, 1965.)

VI.—APPENDIX.

Respondents were asked to put the letters S.A. (Strongly agree), A. (Agree), U. (Undecided), D. (Disagree), or S.D. (Strongly disagree) against the following statements of opinion:

1. If play and activity methods were adopted in my School, examination results would suffer (.64)
2. The social and moral development of the child should be the sphere of the parents NOT the teachers (.62)
3. If the syllabus is based primarily on the children's interests, there will be very little to teach that is really worthwhile (1.26)
4. The only reasonable way of assessing a teacher's worth is through the number of passes he secures in examinations (.70)
5. Children have such strange interests that it is NOT possible to use them as the starting point in lessons (.76)
6. There is plenty of time after obtaining the School Certificate for the child to concentrate on his social development (.96)
7. If classes are streamed on the basis of achievement, class teaching is not only possible, but highly desirable (1.52)
8. Talking in class should NOT be tolerated (1.02)

9. Children should be shamed before their peers for misdemeanours, if high standards are to be maintained (.92)
10. In school we should be concerned not so much with learning as with adjustment to the environment (1.68)
11. The examination system should be modified to cater for the emphasis now being placed on modern teaching techniques (.62)
12. If we pander to a child's likes and dislikes in the learning situation, we shall NOT train him for life (1.16)
13. If he does not know the answer to a question, a skilful teacher can usually talk his way round it (.92)
14. There is nothing like the good old grind to achieve the best results (.90)
15. Facts accepted merely on the teacher's authority are easily forgotten (.72)
16. We must give the child sufficient practice in doing what is unpleasant in school in order to fit him for life (.80)
17. The use of sarcasm is an effective technique for keeping the child in his place (.90)
18. Children cannot be expected to rediscover laws and principles for themselves (.66)
19. The more difficult the subject matter of lessons the better it is for the development of intelligence (.78)
20. Teachers should be benevolent dictators (.94)
21. The imparting of accurate information should be the main concern of the teacher in the classroom (1.26)
22. Individual methods would lead to chaos in my school (.84)
23. The child's mind can be used to store facts, which although they cannot be understood by him at the time, will be of immense use to him in later life (1.12)
24. To make a child learn it is sometimes necessary to resort to corporal punishment (1.38)
25. The class teaching method is the only efficient way of teaching large numbers of children (.90)
26. Chanting in unison is an excellent aid to memory (.76)
27. For effective teaching there should be complete silence with only the teacher doing the talking (.86)
28. If children are to be made proficient in the use of the English language, they should be taught grammatical rules from as early an age as possible (1.1)
29. Teaching facts is NOT the most important part of education (.62)
30. Individual competition is the best incentive to learning (1.0)
31. The teaching methods that were good enough for us should be good enough for the children we teach (.64)
32. Where school discipline is not a major problem, traditional methods of teaching are most effective (.76)
33. In teaching we should define concepts first so that the children will know what is meant (.64)
34. Pandering to children's interests would lead to a lowering of academic standards (1.04)
35. A good teacher does not have to rely on audio-visual aids (.64)
36. There is no limit to the number of facts that children can assimilate provided they are given accurately by the teacher (.88)
37. It should never be necessary to use corporal punishment in the teaching situation (1.12)
38. It is better that the lesson remain unlearned than that the child should be forced to learn it (.78)
39. If teachers became too friendly with the children discipline would suffer (.78)
40. Children have a deep respect for those who rule with an iron hand (.80)

N.B.—Decimal figures in brackets indicate the discriminating indices of the various items.

EFFECTS OF RESTRICTION OF RANGE AND TEST UNRELIABILITY ON CORRELATION BETWEEN MEASURES OF INTELLIGENCE AND CREATIVE THINKING*

BY KAORU YAMAMOTO
(*University of Iowa*)

SUMMARY. The Lorge-Thorndike Intelligence Test, Level 3, Form A, and the Minnesota Tests of Creative Thinking (Torrance, 1962) were given to groups of 461 and 827 fifth-grade American children. Subjects were divided into four groups, according to I.Q. range: below 90, 91 to 110, 111 to 130, and above 130. Linear correlations between I.Q. and creativity score for the whole population and for the sub-groups gave few significant coefficients. Corrections for explicit selection and for unreliability of the creativity measures indicated that the 'true' correlation might be as high as .88. There was consistent decrease in size of the correlation as the level of I.Q. of the sub-groups became higher, this lending support to the idea that, beyond a certain minimum level of intelligence, being more intelligent, does not guarantee a corresponding increase in the creativity. The results do not, however, support the view that creativity is an entity independent of other facets of human intellect.

I.—INTRODUCTION.

STUDIES have reported little relation between measures of intelligence and creative thinking, and it seems generally agreed (Taylor and Holland, 1962) that the correlation is low (.20 to .40) in the general, unselected population, and practically zero in selected (high ability) populations. However, the wide and vigorous reaction to the published work of Getzels and Jackson (1962), Torrance (1962), and others has called this generalization into question. Thus, Burt (1962), de Mille and Merrifield (1962), Thorndike (1963a; 1963b), and Vernon (1964) have challenged the tendency to contrast the two constructs—intelligence and creative thinking, as if they were mutually exclusive. Cureton (1964) pointed out how recognized technical requirements, laid down by the theory of mental measurements, had been ignored by research workers in the construction and application of their novel tests for creativity. Marsh (1964) has re-analyzed the Getzels-Jackson data (1962) and corrected their figures (a) for explicit selection (Gulliksen, 1950, p. 135) of the intelligence variable and (b) for unreliability of the creative thinking measures. Ripple and May (1962) empirically demonstrated the effects of the restricted range of intelligence, and cautioned investigators against accepting the low correlations between intelligence and creativity at their face value.

The present investigation is a partial replication and extension of the Marsh (1964) and Ripple-May (1962) studies with two sets of data obtained from fifth-grade children in two public school systems.

* The author is grateful to the Edina-Morningside (Minnesota) and Cuyahoga Falls (Ohio) school systems for their co-operation in the collection of data. A series of studies involving the Edina-Morningside system was supported by the Bureau of Educational Research, University of Minnesota, while that in the Cuyahoga Falls system was subsidised by the United States Office of Education, Co-operative Research Project No. 2021 (1963-65). Special thanks are due to Dr. Ruth B. Clayton of the Cuyahoga Falls system and Mr. Herbert E. Hartshorn of the Edina-Morningside system for their co-operation.

II.—PROCEDURES.

Subjects.

In school system A, a total of 461 fifth-graders (49.7 per cent. boys) in nineteen classes from six elementary schools during the academic year 1961-62 participated in the study. This system is situated in an exclusive suburb of a large metropolis (1960 Census population, 796,283) in South Eastern Minnesota.

School system B, on the other hand, is situated in a suburb of a middle-sized industrial city (1960 Census population, 290,351) in North-Eastern Ohio; and here 827 fifth-graders (47.3 per cent. boys) in thirty-one classes from ten elementary schools during the academic year 1963-64 constituted the subjects.

Procedures.

In both systems the Lorge-Thorndike Intelligence Test, Verbal Battery, Level 3, Form A, and the Minnesota Tests of Creative Thinking (Torrance, 1962) were administered to pupils in groups. In system A the testing was done in September—October of 1961; and in system B in November, 1963. From the tests of creative thinking (Ask-and-Guess Test and Test of Imagination*) five scores—for Fluency, Adequacy, Flexibility, Originality, and Elaboration respectively—were derived by a scheme which I have suggested in an earlier paper (1962); and the total of these five was used as a single index of creative thinking.

For the total creativity score, interscorer reliabilities of .99 (system A, two scorers over 64 protocols) and .97 to .99 (system B, four scorers over 76 protocols) were obtained. Test-retest reliabilities were .87 in the system A study (twenty-two college subjects over an interval of twelve weeks) and .79 in the system B study (seventy upper-elementary subjects over eight weeks). The validity was established in both cases by reference to teacher nominations of creative pupils (Yamamoto, 1963; Yamamoto, in press).

In each system the subjects were divided into five subgroups according to their I.Q. range: (I) I.Q. 90 or below, (II) I.Q. above 90 but below or equal to 110, (III) I.Q. above 110 but below or equal to 130, (IV) I.Q. of 131 or above, and (V) full-range I.Q. (pooled or unselected group). Within each subgroup and sex nineteen subjects were then randomly selected, and all subsequent analyses were based on these samples.

III.—RESULTS.

Means and standard deviations of I.Q. and creativity scores are shown in Table I for each system, subgroup and sex. Because of the extremely small number of subjects in system A falling in group I, this subgroup was excluded from analysis. On I.Q. no significant sex differences were observed in any subgroup, and in creative thinking the only significant difference was that obtained in system A, subgroup IV where the average for the boys was significantly higher than that of girls. Almost inevitably the selection of pupils to form small groups, graded according to intelligence, has restricted the range and, therefore, the standard deviation of I.Q.s. within each subgroup, but it has not appreciably restricted those of the marks for creativity.

* The Test of Imagination included the Product Improvement (Dog), Unusual Uses (Dog), Unusual Uses (Tin Can), and Circles tasks in system A. All of these, except the Unusual Uses (Tin Can) test, were applied in system B.

TABLE 1

NUMBER OF SUBJECTS BY SEX AND SUBGROUP AND MEANS AND STANDARD DEVIATIONS OF INTELLIGENCE AND CREATIVE THINKING SCORES.†

School System	Subject					I.Q.			Creativity			
	Sub-group	I.Q. Range	Sex	N†	Range	M	S.D.	Range	M	S.D.	Range	
A	I	Below 90 ..	M	3	—	—	—		—	—		
			F	1	—	—	—	—	—			
			M+F	4	—	—	—	—	—			
	II	110 to 90 ..	M	72	19	105.6	4.2		187.5	53.2		
			F	72	19	104.8	4.4		215.1	54.9		
			M+F	144	38	105.2	4.3		201.3	55.8		
	III	130 to 110	M	130	19	121.1	5.5		213.8	58.2		
			F	129	19	118.4	5.0		238.1	58.4		
			M+F	259	38	119.7	5.4		226.0	59.6		
	IV	Above 130 ..	M	24	19	138.3	5.9		256.4	63.6		
			F	30	19	135.1	3.6		209.6	48.1		
			M+F	54	38	136.7	5.1		233.0	61.1		
	Pooled			M	229	19	114.4	12.2	78-150	213.2	73.0	96-498
				F	232	19	116.4	9.6	83-150	205.7	56.4	81-409
				M+F	461	38	115.4	11.0	78-150	209.5	65.3	81-498
B	I	Below 90 ..	M	36	19	84.8	3.1		139.6	42.2		
			F	19	19	85.5	4.7		161.7	48.7		
			M+F	55	38	85.2	4.0		150.6	46.9		
	II	110 to 90 ..	M	185	19	102.5	4.9		155.9	39.6		
			F	184	19	103.8	4.5		181.9	42.3		
			M+F	369	38	103.1	4.8		168.9	43.3		
	III	130 to 110	M	150	19	120.3	5.8		189.5	56.3		
			F	203	19	118.1	4.2		200.0	49.8		
			M+F	353	38	119.2	5.1		194.7	53.4		
	IV	Above 130	M	20	19	136.8	4.6		192.0	37.8		
			F	30	19	137.8	4.8		218.9	57.8		
			M+F	50	38	137.3	4.7		205.4	50.7		
	Pooled			M	391	19	106.4	15.2	58-147	164.3	44.9	52-389
				F	436	19	113.1	12.2	69-146	191.8	43.3	65-372
				M+F	827	38	109.7	14.1	58-147	178.1	46.2	52-389

† N denotes the total number of subjects, and n the number selected for the sub-groups. M denotes mean and S.D. standard deviation.

Correlational data are summarized in Table 2. The table also shows the 'confidence intervals' for the corrected and uncorrected correlations. A confidence coefficient of .95 has been adopted. Thus, for example, the last figures at the bottom right of Table 2 mean that, for boys and girls pooled together, we may safely infer a confidence interval of .26 to .74. In other words, this interval is expected to cover the 'true' value of correlation of creativity with intelligence in 95 per cent. of the time (confidence of .95, or roughly 19 to 1), and the most probable value of this correlation is .54.

First, a linear correlation was computed between I.Q. and creativity scores. It was found that few correlations were significantly different from zero. As might be expected, correlations for subgroups I, II, III, and IV were lower than correlations for subgroup V (unselected). There was a rather consistent tendency for boys to yield a higher correlation than girls, but in no case was the sex difference significant. Accordingly, sexes were combined within each subgroup for further analyses.

TABLE 2
CORRELATIONS BETWEEN MEASURES OF INTELLIGENCE AND CREATIVE THINKING WITH
CORRECTIONS FOR RESTRICTION OF RANGE (a) AND ATTENUATION (b).

School System	Subject		Raw Correlation		Correction (a)		Correction (b)			
	Sub-group	I.Q. Range	Sex	r	CI†	k‡	r	r	CI†	
A	I	Below 90	M	—	—	—	—	—	—	
			F	—						
	II	110 to 90	M+F	—	—	3.72	-.47**	-.53**	-.73 to -.25	
			M	-.16						
	III	130 to 110	F	-.09	-.44 to .19	—	—	—	—	
			M+F	-.14						
	IV	Above 130	M	.34	-.24 to .39	2.95	.23	.26	-.07 to .54	
			F	-.09						
			M+F	-.08						
			M	-.18						
B			F	.13	-.28 to .35	3.13	.14	.16	-.17 to .45	
			M+F	.04						
	Pooled ..		M	.42	.01 to .59	1.45	.45**	.51**	.23 to .71	
	F		.19							
	M+F	.33*								
	I	Below 90	M	.37	-.01 to .57	3.99	.78**	.88**	.78 to .94	
			F	.26						
	II	110 to 90	M+F	.31	-.11 to .51	3.36	.61**	.69**	.48 to .83	
			M	.21						
	III	130 to 110	F	.18	-.40 to .24	3.11	-.27	-.30	-.56 to .02	
M+F			.22							
			M	-.20	-.34 to .30	3.38	-.08	-.09	-.40 to .24	
			F	.12						
	IV	Above 130	M+F	-.09						
			M	-.01						
			F	-.09						
			M+F	-.02						
Pooled ..		M	.45	.08 to .63	1.13	.48**	.54**	.26 to .74		
F	.22									
M+F	.39*									

† CI denotes the confidence interval, see p. 302. * $p < .05$. ** $p < .01$.
‡ k indicates the ratio of the presumable standard deviation for intelligence in the entire population from which the children were drawn (16 I.Q. points on the Lorge-Thorn-dike test) to the calculated standard deviations for several sub-groups.

Next, a correction for explicit selection (restriction of range) was made on a nomograph prepared by Michael, Jones, Gaddis, and Kaiser (1962). It was found that, in spite of the low values for the raw linear correlations (none above .40 for pooled boys and girls), the corrected correlation could be as high as .78, and half the correlations were statistically significant. Finally a correction for the unreliability (attenuation) of the creativity measures was made in the range-corrected correlations. For this purpose the test-retest reliability of .79 obtained from upper-elementary children over an interval of eight weeks (Yamamoto, *in press*) was employed.* The results indicated that the 'true' correlation might be as high as .88 and the confidence interval covers as high a value as .94.

IV.—DISCUSSION.

In general, the critics' objections seem valid. The low correlation between measures of intelligence and creative thinking reported by investigators might result rather from the limited representation of ability range in subject groups and of ignoring the unreliabilities of measuring instruments than from the underlying relation actually existing in the general population. In the present study the correlations corrected for these attenuating sources (without, however, any correction for unreliability of intelligence tests) were above .50 for the total fifth-grade groups, and the confidence intervals at the .95 level covered a range of .25 to .75. Although the two subject groups A and B were different in many respects, the correlational values for the entire range of I.Q. were quite similar (.51 and .54) and corresponded roughly with the corrected values reported by Marsh (1964) for the Getzels-Jackson data (.28 to .51 for five tests of creative thinking; average, .46). Ripple and May (1962), on the other hand, reported correlations of .11 to .73 for nine creativity tests (average, .74) administered to a 'heterogeneous' group of seventh graders with I.Q. ranging from 72 to 133.

Although the subgroups in the present study revealed much variation in the size of the correlations from one system to another (some even having negative signs), there was a consistent decrease in the size of correlation as the I.Q. level of subgroups became higher. In the brighter subgroups (III and IV), the correlations were much lower than in the duller (I and II). In group IV (I.Q. of 131 and above) the coefficient remained quite low even after the corrections. This would seem to support the concept of a 'threshold of intelligence' (Anderson, 1960; Torrance, 1962; Yamamoto, 1964; Yamamoto 1965). In McKinnon's words, "It is clear, however, that above a certain required minimum level of intelligence which varies from field to field and in some instances may be surprisingly low, being more intelligent does not guarantee a corresponding increase in creativeness. It just is not true that the more intelligent person is necessarily the more creative one" (MacKinnon, 1964, p. 311).

Nevertheless, it seems wise for investigators not to overstate their case for creative thinking as if it were an entity independent of any other facets of human intellect. Burt put it succinctly as follows: "... we must, I think, conclude that the weight of the evidence is strongly against the somewhat simplified interpretation ... that there are just 'two basic cognitive or intellectual modes,' the 'creative' and the 'intelligent,' and similarly, two distinct types of 'gifted students' " (Burt, 1962, p. 297). It would appear sensible that,

* If a higher value had been used, the resultant correlation would have naturally been lower.

according to suggestions by such writers as Burt (1962), Cureton (1964), Marsh (1964), Thorndike (1963a; 1963b), and Vernon (1964), we should regard creativity tests as complementary components in new and more inclusive measures of human intellectual behaviour, and not as a measure wholly independent and exclusive of the general factor of intelligence.

V.—REFERENCES.

- ANDERSON, J. E. (1960). The nature of abilities. In E. P. Torrance (Ed.), *Education and talent*. Minneapolis: Univer. of Minnesota Press, pp. 9-31.
- BURT, C. (1962). The psychology of creative ability. *Brit. J. Educ. Psychol.* 32, 292-298.
- CURETON, LOUISE W. (1964). Creativity research and test theory. *Amer. Psychol.* 19, 136-137.
- DE MILLE, R., and MERRIFIELD, P. R. (1962). Book review of *Creativity and Intelligence* by J. W. Getzels and P. W. Jackson. *Educ. and Psychol. Meas.* 22, 803-808.
- GETZELS, J. W., and JACKSON, P. W. (1962). *Creativity and Intelligence*. New York: Wiley.
- GULLIKSEN, H. (1950). *Theory of Mental Tests*. New York: Wiley.
- MACKINNON, D. W. (1964). The nature and nurture of creative talent. In R. E. Ripple (Ed.), *Readings in Learning and Human Abilities*. New York: Harper and Row, pp. 305-323.
- MARSH, R. W. (1964). A statistical re-analysis of Getzels and Jackson's data. *Brit. J. of Educ. Psychol.* 34, 91-93.
- MICHAEL, W. B., JONES, R. A., GADDIS, L. W., and KAISER, H. F. (1962). Abacs for determination of a correlation coefficient corrected for restriction of range. *Psychometrika*, 27, 197-202.
- RIPPLE, R. E., and MAY, F. B. (1962). Caution in comparing creativity and I.Q. *Psychol. Reports*, 10, 229-230.
- TAYLOR, C. W., and HOLLAND, J. L. (1962). Development and application of tests of creativity. *Rev. of Educ. Res.*, 32, 91-102.
- THORNDIKE, R. L. (1963). The measurement of creativity. *Teachers' College Record*, 64, 422-424 (a).
- THORNDIKE, R. L. (1963). Some methodological issues in the study of creativity. In E. F. Gardner (Ed.), *Proceedings of the 1962 Invitational Conference on Testing Problems*. Princeton, New Jersey: Educ. Test. Serv., pp. 40-54 (b).
- TORRANCE, E. P. (1962). *Guiding Creative Talent*. Englewood Cliffs, New Jersey: Prentice Hall.
- VERNON, P. E. (1964). Creativity and intelligence. *Educ. Res.*, 6, 163-169.
- YAMAMOTO, K. (1962). *Revised Scoring Manual for Tests of Creative Thinking: Forms VA and NVA*. Minneapolis: Bur. Educ. Res., Univer. of Minnesota.
- YAMAMOTO, K. (1963). Relationships between creative thinking abilities of teachers and achievement and adjustment of pupils. *J. of Exp. Educ.*, 32, 3-25.
- YAMAMOTO, K. (1964). Threshold of intelligence in academic achievement of highly creative students. *J. of Exp. Educ.*, 32, 401-405.
- YAMAMOTO, K. (1965). Multiple achievement battery and repeated measurements: a postscript to three studies on creative thinking. 16, *Psychol. Reports.*, 367-375.
- YAMAMOTO, K. (1965). *Creative Thinking Abilities and Peer Conformity in Fifth-Grade Children*. (Final report, Co-operative Research Project No. 2021) Kent, Ohio: Kent State Univer., in press.

(Manuscript received 4th January, 1965.)

TEACHERS' PERSONALITY RATINGS OF PUPILS IN SCOTTISH PRIMARY SCHOOLS

BY A. MORRISON, D. MCINTYRE AND J. SUTHERLAND

(*Moray House College of Education, Edinburgh*)

SUMMARY. Thirty-four teachers in primary schools in a Scottish city rated boys and girls in their classes on twenty-five personal characteristics, the characteristics being chosen largely on the criterion of their relevance, in the eyes of teachers, to school behaviour. In addition, measures were taken of the sociometric status, attendance and I.Q. of the pupils. The pupils were members of classes of 11—12-year-olds. Teachers were selected to give a balanced representation for both sexes, married and single women, and age groups; and were chosen from schools from various social backgrounds.

In the first instance, a correlation matrix was computed for each of the sixty-eight single sex groups of pupils. These individual matrices were combined to give six combined matrices—for men, single women and married women, of boys and of girls. The combined matrices were analysed by the method of principal components, each giving six factors which were then rotated by the Varimax method.

Of the six rotated components obtained from each analysis, the first three accounted for between 48 per cent. and 66 per cent. of the total variance. These components show three highly consistent dimensions being used by the teachers in their assessment of pupils, the positive ends of which may be described, respectively, as 'good behaviour,' 'high attainment' and 'sociability and social leadership.'

Results are discussed in relation to differences between the present analyses, to two broad evaluative ratings of the 'pleasant' child and the 'pupil worth taking trouble over,' to teachers' assessments of attainment, and finally, to results obtained from studies of teachers in secondary schools.

I.—INTRODUCTION.

SEVERAL studies have already been published which report on the dimensions used by teachers in English secondary schools in making assessments of the personal attributes of their pupils (Hallworth, 1961, 1962). They have been consistent, both with regard to types of schools and sex of pupils, in demonstrating the use of two main dimensions—the assessment of 'reliability and conscientiousness' to 'social extraversion.' The inclusion, in the more recent studies, of additional measures made by teachers (Hallworth, 1964) or pupils (Hallworth and Morrison, 1964) has not altered this basic picture, although it has made possible the identification of a dimension concerned with attainment, and suggested the existence of a further one on the assessment of 'non-academic leadership.'

Although no such work has hitherto been reported, there would appear to be several arguments in favour of similar investigations at the primary school level, especially if these would make possible comparisons with secondary school results and, at the same time, provide for a more extensive examination of a number of personal and social variables among teachers and schools which may possibly bear upon the teacher's perception of pupils. Some of these arguments relate to matters of general interest concerning the relative importance to teachers of the personal attributes of their pupils and the ways in which these attributes are organised when pupils are assessed, and to the different

situation which primary as compared to secondary school teachers are in regarding the nature of their work, the full-time teaching of one class and the lower age of pupils. Other arguments concern more specific interests such as the increasing use of teachers' estimates of attainments for the purpose of transfer to secondary education.

This article reports on the first stage of an investigation into teachers' assessments of pupils in sixteen Scottish primary schools. The techniques used were similar to those of previous English secondary school studies, but were extended to permit at this stage not only a comparison of assessments made for boys and for girls, but also between categories of male and female teachers.

II.—PLAN OF THE INVESTIGATION.

The subjects were thirty-four teachers and their pupils—some 1,200 boys and girls—in primary schools in a Scottish city. The teachers and their pupils were chosen to satisfy several criteria. To control the possible effects of widely differing ages of pupils upon assessments made by teachers, and, secondly, to ensure that a large enough sample of male teachers would be available, classes of 11—12-year-olds were originally selected. The choice was then further narrowed down to provide a sample of teachers which was equally representative of men, married women and single women, a full range of age of teacher within each of the three categories, and a balanced representation of the three categories of teachers across schools serving middle class, urban and suburban working class, and mixed social class neighbourhoods. For this selection of schools according to social class of neighbourhood the advice was taken of a sociologist who was well acquainted with schools in the city and of practising teachers. It had originally been intended to use twelve teachers from each of the three categories, but in the event it was not possible to obtain twelve married women teachers since relatively few were teaching classes at this age level.

Once the requirements described above had been met a number of the remaining classes of this age group chosen from other schools were used in a number of short pilot studies which preceded the main investigation.

To obtain most of the data used for the present analyses teachers were asked to make ratings of their pupils on a number of personal characteristics. In this respect the technique had much in common with that used in previous secondary school studies and in numerous studies of general personality dimensions in children and adults. However, whilst the rating procedure was the same, the characteristics to be assessed by the teachers were chosen largely on their manifest relevance to the classroom situation. A preliminary list was prepared of characteristics formed the subject of the first pilot study. The choice of characteristics and copies sent to a sample of teachers, similar in general terms that we were interested in the general relevance of these characteristics to the pupils they taught, then were asked to rank them in order of the frequency of occasions on which they found themselves discussing them, and, secondly, to append to the original list any other characteristics which they felt were important but had been omitted. On the basis of the returns (shown in Table 4) a final list was drawn up. Since teachers in the main study were to be asked to make five point bipolar scale ratings of pupils on each of the characteristics it was also necessary to produce brief behaviour descriptions for the extremes of each scale; the suitability of these was then examined as part of the second pilot experiment.

The final list of characteristics to be rated by teachers consisted of twenty-five items, to which were added three further measures—ratings of sociometric status, attendance/absenteeism and I.Q., based on the results of a sociometric questionnaire given to boys and girls separately in each class, class register records of attendance over the previous two years, and verbal reasoning tests carried out as part of the standard procedure for pupils approaching the stage of promotion to secondary education. Regarding this final measure, with some classes it was necessary to base I.Q. ratings on results from a Moray House Picture Test since the standard verbal reasoning test had not yet been given to the pupils. The full list of measures applied to pupils was as follows :

- | | |
|---------------------------------|-------------------------------------|
| 1.—English attainment. | 15.—Calmness/excitability. |
| 2.—Games ability. | 16.—Leadership. |
| 3.—Quietness/talkativeness. | 17.—Home background. |
| 4.—Physical maturity. | 18.—Originality. |
| 5.—Popularity. | 19.—Speech. |
| 6.—Carefulness/carelessness. | 20.—Arithmetic attainment. |
| 7.—Social confidence. | 21.—A Pupil worth taking trouble |
| 8.—Persistence. | over. |
| 9.—Courtesy. | 22.—Co-operation with other pupils. |
| 10.—Enthusiasm/laziness. | 23.—Pleasant pupil. |
| 11.—Gentleness/roughness. | 24.—Attentive/inattentive. |
| 12.—Personal appearance. | 25.—Cheerfulness. |
| 13.—Co-operation with teachers. | 26.—Sociometric status. |
| 14.—Trustworthiness. | 27.—Attendance/absence. |
| | 28.—I.Q. |

III.—ADMINISTRATION OF MEASURES.

Since a large number of schools and of teachers were involved in the investigation it was necessary to send by post to each teacher all the materials and instructions needed for making the ratings. To ensure that all instructions were clear, consistent and comprehensive, a small group of teachers not otherwise involved in the study were asked to give their assistance. Each of these teachers was provided with two sets of ratings blanks—one for girls, the other for boys in their class—and asked to make separate ratings of boys and of girls in accordance with the following set of duplicated instructions :

- (1) One foolscap sheet is provided for ratings on each personality trait. Please complete all ratings on one sheet before proceeding to the next.
- (2) Each trait is described in terms of a pair of opposites corresponding to the two halves of the scale.
- (3) Names of pupils are entered in the left hand column and the rating should be made by means of a tick in one of the five boxes opposite each name.
- (4) The box labelled 3 'average' represents the average for this class of pupils (NOT the average for all pupils).
- (5) Many pupils in this class will be near the average for the class, but some will be above, much above, below and much below average.

- (6) We shall therefore be grateful if you will distribute your ratings on each trait approximately according to the following proportions:

1	2	3	4	5
Much below average 10%	Below Average 20%	average 40%	Above average 20%	Much above average 10%

- (7) We suggest that it is easiest to complete the ratings in the following order: Column 1—10 per cent. much below average, column 5—10 per cent. much above average, column 2—20 per cent. below average, column 4—20 per cent. above average, column 3—remaining 40 per cent.

The teachers who took part in this pilot study were then asked to meet the writers to discuss any difficulties they had encountered either in interpreting the instructions or in making their ratings. It was clear from discussion that, whilst they had no difficulties with the instructions, and had only one or two suggestions to make regarding the description of characteristics, there had been difficulty in accepting the principle of normal distribution of ratings, particularly on a few items. This difficulty was to be encountered again in the main sample of teachers.

In the final stage sets of amended ratings materials and instructions were sent to teachers taking part in the main study, together with a request for records of attendance and I.Q. of pupils. When ratings had been completed and returned, the writers visited each teacher in order to administer a socio-metric questionnaire to the class, and to ask each teacher to complete the Manchester Scales on Opinions about Education (Oliver and Butcher, 1962). Visits also gave an opportunity to ask a minority of teachers to amend some of their ratings distributions to conform more closely to a normal distribution.

IV.—STATISTICAL ANALYSIS.

For each of the single-sex class groups a correlation matrix was computed, using the twenty-eight measures of pupils. The sixty-eight matrices which were obtained—thirty-four for boys, and the same for girls—were then combined, using *z* scores, to give matrices of average correlations in accordance with the following categories of teachers:

- Ratings by married women, single women and men of boys and of girls.
- Ratings by teachers in schools in middle class, urban working class, suburban working class, and mixed social class neighbourhoods. (Separate analyses in each case for boys and for girls.)
- Ratings of boys and of girls by teachers classified into two distinct groups according to age and length of service.
- Ratings of boys and of girls by teachers classified according to high and low scores on the three Manchester Scales of Opinions on Education.

A principal components analysis was then made of each matrix of average correlations. Finally, the principal component axes were rotated by the Varimax method (Kaiser, 1959). Only the results for category (a) above are reported at present.

The writers wish to acknowledge the assistance given by Mr. H. J. Hallworth, University of Birmingham, in the later stages of the analysis of data.

V.—RESULTS.

The results of the Varimax analyses of the data are presented in the following tables, Tables 1 and 2 showing loadings on the three major components and Table 3 a summary of the highest loadings on the remaining components.

Table 4 summarizes the information obtained from individual teachers' rankings, made in the first pilot study, of the importance in their eyes of the characteristics of pupils.

TABLE 1

LOADINGS OF VARIMAX FACTORS OBTAINED FROM COMBINED CORRELATIONS MATRICES FOR RATINGS OF BOYS, BY SINGLE WOMEN, MARRIED WOMEN AND MEN.

	FACTOR I			FACTOR II			FACTOR III		
	S/W	M/W	M	S/W	M/W	M	S/W	M/W	M
1. English Attainment ..	16	25	32	78	84	71	10	21	25
2. Games Ability	06	-08	07	13	30	14	76	55	31
3. Quiet/Talkative	78	63	62	-03	09	-17	-13	01	-26
4. Physical Maturity	00	06	05	13	15	34	12	30	01
5. Popularity	10	43	23	09	16	15	78	43	62
6. Carefulness/Carelessness	57	67	58	58	45	42	18	-13	17
7. Social Confidence	-03	13	05	31	52	31	44	68	64
8. Persistence	63	58	49	57	63	46	27	19	37
9. Courtesy	60	86	78	24	06	21	-07	16	24
10. Enthusiasm/Laziness ..	32	62	41	62	44	54	27	38	40
11. Gentleness/Roughness	70	80	74	14	10	27	-29	-32	-15
12. Personal Appearance	25	57	62	58	26	22	-01	-06	21
13. Co-operation with Teachers	72	76	77	22	29	20	07	35	28
14. Trustworthiness	54	76	72	44	28	36	06	22	26
15. Calmness/Excitability	71	45	62	06	32	-06	20	01	02
16. Leadership	02	21	03	21	44	20	75	53	70
17. Home Background....	30	52	48	56	59	57	16	15	33
18. Originality	08	38	18	60	53	52	22	54	54
19. Speech	26	51	34	51	44	24	03	21	61
20. Arithmetic Attainment	15	24	16	70	85	69	41	15	40
21. A Pupil worth Taking Trouble Over	48	57	41	58	63	48	25	36	41
22. Co-operation with other pupils	12	62	27	23	40	12	53	28	76
23. Pleasant/Unpleasant ..	20	73	58	26	02	15	29	46	52
24. Attentive/Inattentive	68	80	58	49	43	41	13	19	39
25. Cheerfulness	-26	18	00	27	12	11	51	85	84
26. Sociometric Status ..	05	09	00	07	04	09	54	11	15
27. Attendance/Absence	10	04	-08	16	-02	35	14	-02	-01
28. I.Q.	04	08	20	63	79	74	-03	14	05
% of total variance	16.9	27.3	20.3	18.3	19.2	14.7	12.7	12.3	17.7

Decimal points omitted in all factor loadings.

TABLE 2

LOADINGS OF VARIMAX FACTORS OBTAINED FROM COMBINED CORRELATIONS MATRICES FOR RATINGS OF GIRLS, BY SINGLE WOMEN, MARRIED WOMEN AND MEN.

	FACTOR I			FACTOR II			FACTOR III		
	S/W	M/W	M	S/W	M/W	M	S/W	M/W	M
1. English Attainment ..	45	33	35	74	71	74	18	45	37
2. Games Ability	-12	05	03	32	28	38	64	59	62
3. Quiet/Talkative	81	83	59	08	05	11	-37	-31	-33
4. Physical Maturity	01	05	07	10	11	18	11	22	12
5. Popularity	26	35	49	30	36	19	72	62	69
6. Carefulness/Carelessness	73	65	68	28	38	52	07	29	06
7. Social Confidence	10	00	05	52	50	26	65	68	81
8. Persistence	69	52	48	49	61	69	19	34	32
9. Courtesy	85	79	90	17	26	15	20	27	14
10. Enthusiasm/Laziness ..	49	57	51	60	46	56	45	27	41
11. Gentleness/Roughness	83	91	71	-14	-02	-03	-19	-05	-45
12. Personal Appearance	52	58	68	37	28	21	22	46	25
13. Co-operation with Teachers	84	84	81	15	26	24	34	25	27
14. Trustworthiness	79	74	77	35	38	41	25	33	26
15. Calmness/Excitability	65	53	53	30	43	39	-02	-01	-10
16. Leadership	12	21	02	46	31	39	63	69	73
17. Home Background ..	60	52	55	46	52	47	20	41	35
18. Originality	24	28	24	73	73	56	45	45	60
19. Speech	53	42	58	52	51	25	32	46	43
20. Arithmetic Attainment	35	40	44	74	73	70	34	34	34
21. Pupil worth taking trouble over	66	53	60	52	67	53	31	36	41
22. Co-operation with other pupils	44	47	33	05	44	21	72	52	68
23. Pleasant/Unpleasant ..	61	70	70	19	19	25	60	51	49
24. Attentive/Inattentive	82	75	67	38	50	56	11	18	17
25. Cheerfulness	-06	09	11	15	12	11	84	82	83
26. Sociometric Status ..	02	10	03	04	17	03	15	28	12
27. Attendance/Absence	09	06	20	08	03	25	08	-02	08
28. I.Q.	21	12	14	79	84	77	18	10	30
% of total variance	29.2	27.1	26.3	17.9	20.7	17.7	16.8	17.5	19.6

TABLE 3

LOADINGS OVER .30 ON VARIMAX FACTORS IV, V AND VI.

Single Women of Boys	FACTOR			Single Women of Girls	FACTOR		
	IV	V	VI		IV	V	VI
Sociometric status75	.30		Sociometric Status93		
Social Confidence33	.42	.33	Physical Maturity96	
Speech40	.36		Attendance/Absence ..			.97
Cheerfulness40		.43				
Pleasant/Unpleasant ..	.74						
Co-operation with other pupils63						
Courtesy57						
Trustworthiness41						
Enthusiasm/Laziness..	.38						
Co-operation with teachers37						
Pupil worth taking trouble over34						
Gentleness/Roughness..	.31						
Popularity30						
Physical Maturity82					
Leadership36				0	
Attendance/Absence ..			.78				
Men of Boys				Men of Girls			
Games Ability73			Physical maturity89		
Leadership44			Leadership35		
Calmness/Excitability ..	.39			Attendance/Absence ..		.82	
Physical Maturity36		-.67	Calmness/Excitability ..		-.45	
Quiet/Talkative31			Sociometric status95
Sociometric status80					
Popularity32					
Attendance/Absence ..			.71				
Married Women of Boys				Married Women of Girls			
Sociometric status81			Sociometric status78		
Calmness/Excitability ..	.51	.39		Calmness/Excitability ..	.41		
Popularity56			Games Ability36		
Leadership43			Popularity30		
Games Ability38			Physical Maturity89	
Co-operation with other pupils32			Attendance/Absence ..			.96
Physical Maturity71					
Personal Appearance ..		.57					
Speech40					
Quiet/Talkative		-.45					
Attendance/Absence ..			.89				

TABLE 4

FIRST PILOT STUDY—TEACHERS' RANKINGS OF PERSONALITY CHARACTERISTICS OF PUPILS
(56 TEACHERS).

Characteristics	Totals of Ranks Awarded by All Teachers	Ranking in Frequency of occasions on which Characteristic Discussed by Teachers
General Ability	231	1
Carelessness	343	2
Laziness	463	3
Talkativeness	485	4
Co-operativeness with Teacher	552	5
Persistence	638	6
Courtesy	642	7
Ability to use language	672	8
Originality	717	9
Excitability (1)	737	10
Attractiveness of Personal Appearance	770	11
Cheerfulness	771	12
Co-operativeness with other Pupils	779	13
Impertinence	782	14
Pleasantness of Speech.....	789	15
Leadership	790	16
Games Ability (2)	807	17
Social Confidence	810	18=
Sociability	810	18=
Roughness	837	20
Physical Maturity	841	21
Popularity	882	22
Perfect Suitability	943	23

(1) 'Excitability' ranked fifth by single women alone.

(2) 'Games Ability' ranked tenth by men alone.

The most frequently mentioned characteristic added to this list by teachers was the quality of the pupil's home background.

VI.—DISCUSSION OF RESULTS.

The Three Major Dimensions.

The general nature of the three main components is closely similar in all six analyses. On the first component, the highest loadings in nearly all cases are on Courtesy, Co-operation with Teachers and Gentleness, while Trustworthiness, Quietness, Attentiveness and Carefulness also have consistently high loadings. This component may, therefore, be said to define a dimension of the extent to which a pupil's social behaviour conforms with the standards of the teacher.

On the second component the three highest loadings in all cases, except by a slight margin in one, are on I.Q., English Attainment and Arithmetic Attainment; other traits which always have high loadings are Originality and, to a lesser extent, Persistence. The dimension here defined is clearly one of general attainment.

On the third component Cheerfulness has the highest loading in five of the six cases, and Leadership, Social Confidence and Popularity have high loadings in all cases, as do Games Ability and Co-operation with Other Pupils in all but one case. Opposed to these are generally small loadings on Gentleness and Quietness, in no case over .50. The description of this third dimension is more difficult. Previous studies of secondary school teachers' ratings have found a factor similar to this one, generally described as one of Social Extraversion/Introversion, but such a description does not seem entirely adequate in this case for two reasons. Firstly, the traits Quietness and Gentleness are much more closely related to the orthogonal dimension of 'good behaviour' than they are to the negative pole of this one; and secondly, while Cheerfulness and Social Confidence may be described as extravert traits, the other defining traits do not fit so easily under such a heading. A more satisfactory description of this dimension might be 'Sociability and Social Leadership.'

The six analyses can initially be compared in two ways. Firstly, what proportion of the total variance do the first three components extract between them in each analysis? Secondly, to what extent are these components, in each analysis, independent of each other, or, conversely, to what extent are they merely different aspects of a generalised scale of assessment?

The three components account for a considerably larger amount of the variance in the analyses of the ratings of girls than they do in those of boys. There are negligible differences in this respect between the three groups of teachers so far as girls are concerned. For boys, on the other hand, there are sizeable differences, with the largest amount of variance accounted for by these three components being in the analysis of assessments made by the married women, and the least amount being in the single women's analysis.

The independence of any one of the three components may be gauged from the extent to which its pattern of loadings differs from the loadings pattern of each of the others. Judged in this way, the components in the analyses of ratings of boys tend to be more independent than in those of girls; and, comparing the three groups of teachers, the components in the analyses of single women's assessments are the most independent, and the components in those of married women the least.

The analysis in which these three components extract the least amount of variance, and the one in which the least generalised assessments are made, is that of single women's ratings of boys. This analysis is also the one which differs most from the others in the loadings of particular items on these components. The major difference is that several variables, in particular, Pleasantness, Courtesy, Home Background, Enthusiasm, Speech and Personal Appearance, have much smaller loadings on the first component than in the other cases. On the third component, Pleasantness and Originality both have much smaller loadings than in the other analyses.

Further Components.

The loadings on each of the three remaining components extracted for each analysis are in almost every case dominated by one particular variable. Thus the general pattern is for Component IV to be one of Sociometric Status, Component V one of Physical Maturity, and Component VI one of Attendance/Absenteeism. With one or two exceptions, loadings of other variables on these components are small.

Two of the subsidiary components, however, break this general pattern. Component IV for single women's ratings of boys, extracting as it does almost 10 per cent of the variance, is much larger than any of the other fourth components, although, like them, its highest loading is on Sociometric Status. The difference is that in this case there are several other variables with high loadings; indeed, one can consider this component as defining a fourth dimension for single women's ratings of boys, a dimension which does not appear in the other analyses. Together with Sociometric Status, the variables with highest loadings are Pleasantness, Co-operation with other Pupils and Courtesy. While many of the traits with loadings on this component are those which have higher loadings in the 'good behaviour' component in the other analyses, several others are characteristics which are generally associated with the 'sociability and social leadership' dimension. The dimension defined by this fourth component may perhaps be best described as one of the extent to which a boy is judged to be pleasant both by the teacher and by other pupils.

It is interesting to compare the component just described with component IV for married women's ratings of boys, on which the loading for Pleasantness is negligible, but that for Popularity is relatively high. Thus, while married women have a much closer idea as to which boys are popular, they do not themselves find these boys particularly pleasant; on the other hand, single women are relatively ignorant about which boys are popular, and yet those boys who are in fact most or least popular tend to be those whom the teachers themselves assess as, respectively, most or least pleasant. In the four other analyses, however, there is a component in which the Sociometric Status measure is the only variable with a loading of any size, thus indicating that not only do most of the teachers have little idea as to which pupils are popular with their classroom peers, but also that the actual popularity of a child bears hardly any relation to the teacher's general perception of him. This result is perhaps more easily understood in the light of the results of the pilot study which indicate that the extent to which a child is popular is a subject holding relatively little interest for the teachers.

The second exception among the minor components is component IV for men's ratings of boys, which is dominated by a high loading for Games Ability. Whereas in the other five cases Games Ability appears as one of the dominant loadings on Component III, indicating that Games Ability is treated as one aspect of the 'sociability and social leadership' dimension, the men, in rating boys, evaluate Games Ability on a distinct dimension. The minor loadings on this component show a tendency for the men to perceive the boys who are good at games as quiet, calm, physically mature leaders—leaders of a different nature from those assessed highly on the 'sociability and social leadership' dimension. The greater part which men usually play in teaching games to boys in primary schools—a concern which is reflected in the detailed analysis of the rankings made in the first pilot study—clearly influences them to use an assessment dimension which reflects this specific concern.

With regard to the two other components generally found in the analyses, Physical Maturity, especially with girls, is assessed by teachers in almost entire independence of other pupil characteristics. Since results of the first pilot study show that physical maturity is a characteristic of pupils with which teachers are very little concerned, the present results might be explained by its being so unimportant as not to be related to any of the three fundamental stereotyped dimensions commonly used in assessments.

With even higher consistency, the objective assessment of attendance at school is unrelated to the teachers' ratings. For these teachers, attendance/absenteeism is apparently not a factor which affects their perceptions of pupils or their assessments of the home background.

The high negative loadings for Physical Maturity on factor VI of Men's ratings of boys, and for Quietness on factor V of Married Women's ratings of boys do not correspond to actual correlations of any size. The explanation for these results could presumably be found in subsequent residual factors, and must be seen as a consequence of the mathematical techniques; without further evidence they cannot be interpreted as having any significance.

Aspects of Classroom Behavior.

Several pupil characteristics which are particularly related to classroom behaviour, and which teachers in the first pilot study rated as being of particular concern to them, were included in this study. Five of these were Quietness/Talkativeness, Carefulness/Carelessness, Enthusiasm/Laziness, Persistence and Home Background. All these have high loadings on the 'good behaviour' component in each case. Enthusiasm, Home Background and Persistence also have high loadings, and Carefulness moderately high loadings, on the dimension of attainment. Attentiveness, a characteristic not included in the pilot study, has very similar loadings to those of Carefulness. Calmness is another item which has high loadings on the 'good behaviour' dimension, especially in the two analyses of ratings made by single women, for whom it is, by pilot study results, a much more important characteristic than it is for the other categories of teachers. In ratings of girls, but not in those of boys, Calmness also has a moderately high loading on the attainment dimension. It is noteworthy that, of the twenty-three traits presented in the pilot study, those six which appeared as being of the greatest concern, after general ability, for teachers, all have consistently high loadings on the 'good behaviour' dimension. In comparison, those six traits which define the dimension of 'sociability and social leadership' had an average rank of 16. Thus, while teachers make use of both dimensions in their assessments of pupils, it is clear that the 'good behaviour' dimensions is of much the greater importance to them.

Originality, another trait in which teachers showed considerable interest, has a quite distinctive pattern of factor loadings. It is very closely associated with the attainment ratings and with I.Q., but it also, with the exception of single women's ratings of boys, has high loadings on the 'sociability and social leadership' dimension. Furthermore, it is the only trait which is clearly associated with high attainment ratings and yet is not a characteristic of the 'well behaved' pupil.

Speech is a characteristic which also merits attention, for it is here that factor loadings vary most. For ratings of girls it tends to have moderately high loadings on all three components. With boys, however, it is high on the 'good behaviour' component for married women's ratings, high on the attainment component for those of single women, and high on the 'sociability and social leadership' component for those of men. Thus, the clarity and pleasantness of a boy's speech tends to have a different implication for different teachers, according to sex and marital status.

In two cases, a pair of traits have consistently similar patterns of loadings; co-operation with teachers is paired with Courtesy, and Social confidence with Leadership. In effect, traits within these pairs are largely synonymous in meaning.

The Teachers' Value Judgments.

In assessing pupils on many of the characteristics, especially those with high loadings on the first component, the teachers are making implied value judgments. They also rated their pupils, however, on two straightforward value judgments. 'A pupil worth taking trouble over' has, without exception, high loadings on both the dimensions of 'good behaviour' and 'high attainment.' It is closely associated with the variables of Persistence, Enthusiasm and Home Background. Pleasantness, with the exception of single women's ratings of boys in which it is closely associated with sociometric status, has high loadings on the 'good behaviour' dimension, and slightly less high loadings on the 'sociability and social leadership' dimension.

In all the ratings of girls, and in the married women's ratings of boys, over 80 per cent. of the item variance for 'A pupil worth taking trouble over,' and between 74 per cent. and 80 per cent. of the item variance for Pleasantness, is accounted for by the three major dimensions. Apart from single women's ratings of boys for Pleasantness, on which only 19 per cent. was extracted, the item variance accounted for in the other cases was about 60 per cent. Thus, especially for girls, and particularly in the case of married women, judgments as to the relative value of teaching a child and of his/her pleasantness are made very largely in terms of the three main dimensions.

Teachers' Assessments of Attainments.

One of our particular interests was the implications which teachers' perceptions of their pupils had for their assessments of attainment. Together with I.Q., English and Arithmetic Attainment ratings are the defining variables of one of the three major dimensions. English attainment has higher loadings for girls than for boys on the 'good behaviour' and 'sociability and social leadership' components, and Arithmetic attainment has considerably higher loadings for the girls on the 'good behaviour' component than for the boys. There are no systematic differences between the teachers in this respect.

All the ratings of attainment are quite clearly related to I.Q., although in all cases except men's ratings of boys, English and Arithmetic Attainment ratings have noticeably higher loadings on the 'good behaviour' component than do the ratings of I.Q. Also, the attainments loadings on the 'sociability and social leadership' component are higher than those of I.Q. for married women's ratings of girls and for men's ratings of boys. Arithmetic Attainment loadings on this component, but not those of English Attainment, are higher than I.Q. loadings in two analyses—single women's ratings of boys and girls.

Of the traits which have consistently high loadings on the dimension of attainment, all have high loadings also on one of the other dimensions—Originality on the 'sociability and social leadership' dimension, and Enthusiasm, Persistence and Home Background on the 'good behaviour' dimension. There are considerable variations, according to the sex of pupils and the category of teacher, in the loadings of particular traits on the attainment dimension. Of the six analyses, that for men's ratings of boys is the one on which traits other than attainments have generally the smallest loadings, and which accounts for the least variance. With regard to particular traits, Speech has higher loadings for women's than for men's ratings; Social Confidence, Leadership and Co-operation with Other Pupils have high loadings for the ratings made by married women, as do the first two of these for the ratings of girls by single women. An interesting peculiarity is the loading of Personal Appearance for single women's ratings of boys.

It is clear from these results that, while general attainment is perceived as a distinctive dimension of pupils' personalities, and while assessments of it are closely related to objective measures of I.Q. that these assessments are to some extent affected by the teachers' perception of other characteristics of their pupils. This tendency would seem to be greatest for married women and least for men. The association of high attainment with good behaviour might have been expected, but it is interesting to discover that teachers also relate high attainment to traits of leadership and extraversion, especially when rating children of their own sex. The tendency to assess the attainment of girls as more related to other aspects of personality than is the case for boys is consistent with the general tendency to assess girls in a more global way.

Major differences according to sex of pupils and category of teacher.

Between the two sexes there are several consistent differences. In the ratings of all three groups of teachers, the variance extracted on the three major dimensions is greater for girls than for boys, the value judgments made of girls are made more in terms of these dimensions than those of boys, and a more generalised assessment is made of girls than of boys. Thus girls tend to be assessed in a less analytic way than boys. In accordance with the general tendency, attainment ratings for the girls are more closely related to other traits. In particular, teachers associate arithmetic attainment more with good behaviour in the case of girls than they do with boys.

Of the three categories of teacher, the married women's perceptions of their pupils are the least analytic. The men, too, tend to be less analytic in their perceptions than the single women. On the other hand, the men assess the attainments of pupils more independently of other aspects of their behaviour than does either group of women.

Comparisons with Results obtained from Secondary Schools.

In general, the results obtained in this study indicate that these primary school teachers used major dimensions of assessment similar to those reported from studies carried out in secondary schools. In both categories of schools the first three dimensions are of 'good behaviour' (Usually described as 'Reliability and Conscientiousness' in secondary school studies), 'Attainment and I.Q.' and 'social extraversion.' A more detailed comparison is difficult, however, since trait lists across studies are not identical, and, also, secondary school studies have not differentiated so fully on categories of teachers. When these two facts are taken into consideration it becomes clear that the naming of certain dimensions needs to be changed to take account of both additional measures and of significant modifications in loadings patterns—this seems particularly necessary in the case of 'social extraversion'—and that we have to recognise important differences between categories of primary teachers in their use of minor dimensions. Some of the latter differences probably apply to secondary teachers, but, as yet, we do not have strictly comparable categories of teachers on which to confirm this.

Clearly we can trace back to primary school level two important features of secondary teachers' assessments: the use of dimensions rather than independent assessments of discrete traits, and the consistent use of three such dimensions, irrespective of type of school and category of teacher.

VII.—CONCLUSIONS.

From the results obtained the following conclusions can be reached :

- (a) Teachers belonging to the three categories examined use the same three major dimensions for the assessment of pupils. Three exceptions are a 'games ability' dimension for men of boys, for single women of boys, a dimension of sociometric status, pleasantness and several other traits, and, for married women of boys, a dimension on which assessments of popularity are associated with sociometric status.
- (b) The three major dimensions are those of (in positive terms) 'good behaviour', 'high attainment,' and 'sociability and social leadership.'
- (c) The results indicate the use of assessment dimensions similar to those reported from studies of teachers in English secondary schools.
- (d) In general, teachers' assessments of physical maturity, and objective measures of attendance/absenteesim and sociometric status are independent of the assessment of other traits.
- (e) All categories of teachers tend to make a more general evaluation of girls than of boys.
- (f) Assessments of attainment, together with I.Q., are largely carried out on a separate dimension ; however, with variations across the analyses, these assessments are also related to various traits, notably originality and persistence.

ACKNOWLEDGEMENTS.—The writers wish to thank all members of the local Education Authority who were involved in the research, and the Research Committee of Moray House College for the grant which made it possible to carry out the investigation.

VIII.—REFERENCES.

- HALLWORTH, H. J. (1961). Teachers' personality ratings of high school pupils. *J. Educ. Psychol.*, 52, 297-302.
- HALLWORTH, H. J. (1962). A teacher's perceptions of his pupils. *Educ. Review.*, 14, 124-133.
- HALLWORTH, H. J. (1964). Personality ratings of adolescents : a study in a comprehensive school. *Brit. J. Educ. Psychol.*, 34, 171-177.
- HALLWORTH, H. J., and MORRISON, A. (1964). A comparison of peer and teacher personality ratings of pupils in a secondary modern school. *Brit. J. Educ. Psychol.*, 34, 285-291.
- KAISER, H. F. (1959). Computer program for Varimax rotation in factor analysis. *Educ. Psychol. Meas.*, 19, 413-420.
- OLIVER, R. A. C., and BUTCHER, H. J. (1962). Teachers' attitudes to education. *Brit. J. Soc. Clin. Psychol.*, 1, 56-69.

(Manuscript received 4th January, 1965.)

LINEAR AND SKIP-BRANCHING PROGRAMMES: A COMPARISON STUDY*

By JAMES HARTLEY

(*Department of Psychology, the University of Keele*)

SUMMARY. The results of a comparison between a skip-branching and a linear programme designed to teach logarithms to secondary modern school pupils, indicated an initial superiority for the branching programme which disappeared on retest, two weeks later. The test and retest results of the branching programme correlated significantly with intelligence; those of the linear programme did not. Analysis showed that the branching programme catered as well for the lower ability pupils as did the linear one, but that higher ability pupils profited more from the branching than from the linear programme. There was no difference between the times taken to work through the programmes, but there was in error-rate. Attitudes towards the two programmes were similar: at first they were highly favourable, but later, much more neutral.

I.—INTRODUCTION.

PRIOR research^{1, 2, 3, 4, 5, 6}, comparing multiple-choice branching with linear programmes has indicated that branching programmes can teach as effectively as, but often more quickly than linear ones, especially with adults. This conclusion, however, is limited by the paucity of studies reported, restrictions imposed by the subject matter programmed, and the age and ability levels of students used. The purpose of the following experiment was to see how far any of the generalisations made concerning the effectiveness of the multiple-choice branching system (e.g., that it caters better than the linear system for individual differences) could be made for an alternative system called 'skip-branching.' These different types of programme have been described elsewhere^{7, 8}.

Experiments with programmes of the 'skip' kind have provided equivocal findings: Coulson and Silberman¹ found a difference suggesting that branching was superior; Campbell⁹, after several attempts, failed to produce a branching programme that was significantly better than a linear one; Glaser and Reynolds¹⁰ found no significant differences, Knapper's¹¹ results were inconclusive. The total picture of 'linear versus branching' studies is thus confused, and the generalisations drawn from them, limited. The experiment reported below is seen by the writer as one of many necessary before worthwhile conclusions can be drawn.

II.—PLAN OF THE INVESTIGATION.

(i) *Subjects and Materials.*

Sixty-eight secondary-modern school girls, of average age 12 years 9 months, took part in this experiment. Their mean score on Heim's AH4 intelligence test was 61 (range 38-81) and on Vernon's Graded Arithmetic Mathematics test it was 37 (i.e., 13-14 years) (range 27-51). Their mean I.Q. and A.Q. scores (from 1962 11+ data) were 104 (range 94-116) and 110 (range 99-130), respectively.

* This experiment was carried out whilst the author was a recipient of a D.S.I.R. research studentship at the University of Sheffield. It is reported in considerably more detail in HARTLEY, J. (1964). *A Study in Programmed Learning*. Unpublished Ph.D. thesis, University of Sheffield.

The subjects worked with lessons 1 to 13 (except lesson 11) of the writer's linear programme *Logarithms*¹²—each lesson being presented separately in booklet form—and with the parallel lessons of a skip-branching programme also devised by the writer and also in booklet form.

Students were requested to write down in separate notebooks their answers to both the linear and the branching programmes, the essential difference being that if the student made a mistake in the branching programme he was directed through a linear remedial sub-sequence on the left-hand page before he continued to the next main frame. If the student was correct on the main sequence frame he skipped the sub-sequence and proceeded directly to the next main frame.

Both versions of the programme used in this experiment had previously been evaluated, and revised twice. The final test used in this experiment had an internal consistency coefficient of 0.88 (found by using the modified Kuder-Richardson formula : Guildford¹³) and a test retest coefficient (two weeks delay) of 0.60.

(ii) *Experimental procedure.*

Two matched groups were formed on the basis of their scores on the AH4 and Vernon's Mathematical Ability test, both of which were administered two weeks before the experiment began. No significant differences were found on the other measures, A.Q., I.Q. or age.

The sixty-eight subjects were taken from 'A' and 'B' streams (which, unfortunately, had different mathematics teachers). Both forms had mathematics during the same double-period on one day of the week, so it was therefore possible for the branching group to work in one room, and the linear group in another. The classrooms had double desks so pupils sat in pairs. The experimenter alternated between the two classrooms, and a teacher was also present in each. The experimenter, however, spent more time (approximately twice as much) with the branching group, as with this group the teacher was often absent. In the first session the opening instructions for both types of programme were read aloud to the respective groups by the experimenter. It was explained that both books taught the same material, that they used different methods, and that a comparison was being made between them. It was indicated to the subjects that it was very easy to cheat, and they were asked not to do so: it was pointed out that cheating would not only invalidate the experiment, but also that it was in their own interest not to cheat. It was explained that pupils worked on their own, and only asked the experimenter or teacher for help in cases of real difficulty.

The subjects worked for one 'double-maths' period at a time (approximately one hour), and proceeded to the next lesson in the series as they required it. In order to keep the groups as close together as possible, and to avoid time-tabling difficulties at the end of the experiment, subjects were required to reach certain points in the programme in each period. Pupils who had not reached the required point by the end of the period completed up to this point for home-work: this was collected after four days.

To reduce alarm caused by the concept of 'test' the final test was incorporated as part of the programme (i.e., as problems to complete without answers). In the last session pupils were required to finish immediately before these problems, and thus the immediate test was, in reality, given one week later. To measure retention the same test was readministered in a more formal test situation after a further two weeks: the two classes were not, on this occasion, separated

into the linear and the branching groups. Subjects did not work on logarithms between the tests and the results on the immediate test were not communicated to them.

Attitude questionnaires were administered after four and thirteen weeks. The latter period, however, included a four week vacation.

III.—RESULTS.

The main results of the experiment are summarised in Tables 1-3. Subjects who had not completed the necessary sections of the programme were excluded from the analysis. (Unless otherwise stated, the test of significance used was the Mann-Whitney U test: Siegel¹⁴).

(i) *Test and retention results.* Table 1 shows the test and (in brackets) the retest scores of the two groups. The branching groups produced significantly better scores on immediate testing, but this overall significance was not retained after a period of two weeks. Table 2 shows that the superiority of the branching programme showed itself in the higher-ability stream, 2A; the results of this group remained superior to those of the linear group in both the immediate and the retention test (although the level of significance was reduced). The main contribution to the overall decline in significance of the branching programme came from 2B, whose retention scores, unlike those of 2A, were significantly lower than their immediate test scores ($U = 44$ $p < .02$).

TABLE 1

THE MEAN IMMEDIATE TEST AND RETENTION TEST SCORES.
(Retention Test Scores in brackets: test marked out of 46.)

	Linear	Branching	Significance of the difference between means	(Significance of the difference between means)
\bar{x}	33.5 (34.4)	39.1 (36.9)	$p < .01$	(NS)
SD	8.0 (6.2)	4.6 (5.8)		
N	28 (27)	31 (26)		

TABLE 2

THE MEAN TEST AND RETENTION SCORES OF 2A AND 2B (RETENTION SCORES IN BRACKETS).

	Linear	Branching	Significance of the difference between means	(Significance of the difference between means)
2A \bar{x}	34.8 (36.4)	39.5 (40.3)	$p < .02$	$(p < .05)$
SD	4.1 (4.4)	5.2 (4.2)		
N	17 (17)	15 (13)		
2B \bar{x}	31.5 (31.0)	38.8 (33.5)	NS	(NS)
SD	12.4 (7.2)	3.4 (5.2)		
N	11 (10)	16 (13)		
Significance of the difference between means ..	NS (NS)	NS ($p < .002$)		

(ii) *Time taken.* No significant difference was found between the times taken on the two programmes: the linear group averaged 39 frames an hour (s.d.7), and the branching group 20 main sequence frames. This, when expressed in terms of its linear equivalent, was approximately 41 (s.d.7) frames per hour. 2B did not work as quickly as 2A using either programme, but in both cases the differences were not significant.

(iii) *Error rate.* There were significant differences in the number of errors produced by both programmes. The mean error rate per frame for the linear programme was 6.25 per cent. and (on the main sequence frames) for the branching programme it was 11.25 per cent. The distributions, however, were positively skewed, and in addition, for the branching programme there was a steady decline in errors, whereas for the linear programme the reverse occurred. 2B made fewer errors than 2A on both programmes, but this difference was slight.

TABLE 3

INTERCORRELATIONS BETWEEN THE MATCHING VARIABLES AND THE EXPERIMENTAL RESULTS. (Pearson's Product Moment).

	AH.4	11+I.Q.	A-M	11+A.Q.	Time	Error Rate	Test
Linear:							
I.Q.	+0.59†	—	—	—	—	—	—
A-M	+0.17	-0.21	—	—	—	—	—
A.Q.	+0.43*	+0.10	+0.52†	—	—	—	—
Time	+0.21	+0.07	+0.14	+0.27	—	—	—
Error-Rate ..	+0.10	+0.05	+0.18	+0.05	-0.22	—	—
Test	+0.03	-0.04	+0.12	+0.25	-0.08	-0.01	—
Retest	+0.30	+0.13	+0.40*	+0.53†	+0.51†	+0.21	+0.60†
(N=24-32):							
Branching:							
I.Q.	+0.71†	—	—	—	—	—	—
A-M	+0.25	+0.03	—	—	—	—	—
A.Q.	+0.47*	+0.11	+0.62†	—	—	—	—
Time	+0.14	+0.24	-0.11	+0.11	—	—	—
Error-Rate ..	-0.19	-0.21	+0.18	-0.20	-0.20	—	—
Test	+0.30	+0.27	-0.29	-0.09	-0.01	+0.28	—
Retest	+0.52†	+0.52†	+0.43*	+0.35	+0.16	-0.09	+0.58†
(N=22-33):							

* $p < .05$.

† $p < .01$.

(iv) *Intercorrelations.* Table 3 gives the product-moment correlation coefficients found between the matching variables and the experimental results. This table shows that the immediate test results correlated more highly with intelligence for the branching than for the linear programme, and that the retest results from the branching group correlated significantly with intelligence (as measured by both the 11+ data and the more recently administered AH4), whereas those from the linear group did not.

The retest results for both groups correlated significantly with mathematical ability. Time taken and retest scores correlated significantly for the linear but not for the branching group. The remaining correlations of the experimental results with the matching variables (except for test-retest) were not statistically significant.

(v) *Attitudes.* Attitudes towards the programmes were measured after four and thirteen (or effectively nine) weeks. A questionnaire consisting of five statements (e.g., "I enjoyed using this programme") was given. Pupils then had to indicate their position by assenting to one of five points on a scale between two extremes. A further question (question 6) asked them to rank the lessons held that week (including the programmed one) in order of preference. After four weeks there were no significant differences between the groups, or between the streams. Over 60 per cent. of the subjects enjoyed using the programmes, 70 per cent. found them neither hard nor easy, and the remaining 30 per cent. found them easy: 90 per cent. of the subjects found the programme interesting, and 60 per cent. wanted to continue using the booklets. Cheating was admitted (in response to "How often did you look at the correct answer before writing your own?") by a higher percentage of the branching group ("Sometimes, i.e., once every three lessons": linear 7 per cent. branching 21 per cent.) Although no significant differences emerged, the general trend favoured the linear programme. Responses to the sixth question indicated that—after four weeks—the logarithms lesson was ranked as preferable to the conventional mathematics lesson, but not significantly so.

After nine weeks the attitudes of the two groups towards the programme were again not significantly different. There was, however, a significant decline in the favourableness of the attitudes previously expressed. Only 35 per cent. now enjoyed using the programme, only 55 per cent. found it interesting, and only 35 per cent. wanted to continue using the programme. The ranking of mathematics remained constant, but the mean ranking of logarithms was now lower. The mathematics lesson was now ranked significantly preferable to the lesson in logarithms (taken overall). Examination of the responses showed that a large measure of the decline in the popularity of the programmes occurred with 2B, particularly in this respect, with the linear programme. Overall, however, there was again a slight indication of preference for the linear programme: this preference was more marked in the high ability group 2A.

IV.—DISCUSSION OF RESULTS.

A number of suggestions can be made to account for the results shown in Tables 1 and 2, but examination of the data shows some of these suggestions to be suspect, and some of them to be untenable. This examination, for example, did not support (i) the suggestion that the loss of five subjects from the branching group between test and retest could account for the decline in the branching group scores, or (ii) that the differences in test and retest results between the two groups were due to mathematical errors rather than errors in finding and using logarithms, or (iii) that one programme taught one process much better (or worse) than the other. An analysis based on that proposed by Markle¹⁵ showed that there were weaknesses in the programmes, but these were parallel, rather than divergent. (A closer inspection did suggest, however, that the use of antilogarithms was not as well taught by the linear as by the branching programme).

One further suggestion which received some support was that subjects using the branching programme might have made less responses than their counterparts using the linear one, and that this might be more important for the lower ability group. By calculating the average number of sub-sequence frames used by the branching group, and the average number of responses required by these frames, an approximate total of 700 responses per subject is reached. This figure represents a reduction of 25 per cent. on the total number

of responses required by the linear programme. The fact that users of the branching programme thus made a smaller number of responses might well account for the decline in the retention of the lower ability group, particularly as the responses deleted to form this version of the programme were mainly repetitive items.

It is, however, important to note that although the scores of the lower ability group 2B with the branching programme declined, they were still higher than those who had used the linear programme (although not significantly), and, furthermore, that the retention scores of the higher-ability group 2A were better for the branching group than they were for the linear. It is possible to conclude, therefore, that the higher ability students profited from the branching programme more than they did from the linear programme, and that the lower ability pupils were not held any further back than were their counterparts using the linear programme.

This conclusion is further supported by intercorrelations between the matching variables and the experimental results. It would be tempting to explain the fact that the results from the branching group correlated with intelligence—whereas the results from the linear group did not do so—in accordance with previous literature, by stating that the linear programme catered better for individual differences by bringing all subjects to the same level despite variations in initial ability, whereas the branching programme failed to do this.*

Such an explanation would, however, be a misinterpretation of these results. By taking from each group the top and bottom ten subjects on the AH4, and by comparing their results, it can again be seen (Table 4) that the branching programme catered as well as the linear one for the lower intelligence groups, but that the more intelligent pupils profited more from the branching programme than did their counterparts from the linear one. With the retest scores the difference between the top ten subjects of the two groups just escaped the accepted significance level of $p = .05$: nevertheless, the top ten subjects with the branching programme were significantly superior to the bottom ten, who in turn, were not significantly different from the bottom ten using the linear programme.

TABLE 4

THE MEAN TEST SCORES OF THE TOP AND BOTTOM TEN SUBJECTS ON THE AH4.

		Linear	Branching	Significance of the difference between means
AH4 Top 10	\bar{x} S.D.	34.0 8.4	41.9 3.4	$p < .02$
AH4 Bottom 10 ..	\bar{x} S.D.	31.2 10.2	36.7 5.0	NS
Significance of the difference between means ..		NS	$p < .02$	

*Compare, for example, Stolurow's¹⁶ survey of results from linear programmes with the three studies reported by Cavanagh¹⁷.

The overall picture, and the intercorrelations, thus together support the conclusion that, in this experiment, the teaching system which was more immediately sensitive to individuals—the skip branching programme—was more effective than was the system that was based upon previously obtained results, i.e., the linear programme.

In this paper it has not been possible to report in full detail the times taken and the error rates of the different groups (but they are available elsewhere¹⁸). There are two main reasons for this: firstly, to do so would be a complex business involving a discussion of the different variances obtained over different sections of the programme, overall, and by different groups: secondly, the conclusions reached are in a sense suspect because a text-book rather than a machine presentation of the material was used. Before considering the implications of this last point in more detail, a brief summary of the main conclusions of the study of the time data can first be made. These were: (i) that the branching group worked more quickly than the linear group over revision sections in the programme, and (ii) that the mean times taken by both groups per session were roughly equal, but at the close of the experiment the variance was significantly greater with the linear programme, whereas at the beginning of the experiment, this had been the case for the branching programme. The distribution of error rates complemented this picture. Such results indicate the usefulness of time (and error) data in programme evaluation, although it should be pointed out that in this experiment there was no significant correlation between time taken and error rate for either programme. Brooks¹⁹ and Coulson and Silberman¹ have indicated that response latency is a more sensitive measure of difficulty than is error rate.

One disturbing finding concerning the time data was the significant retest correlation with time for the linear programme: the writer would like to treat this as an anomaly, but it has occurred before in pilot studies using this programme. The exact significance of this finding is unknown.

As stated above, conclusions about the results of time and error measures are suspect because of the conditions of the experiment. It was impossible to assess how far the experimental conditions—particularly of finishing off uncompleted work for homework—affected the rate of progress: it is possible that the more able pupils adjusted their rate of progress in order to complete the amount of work set in class, but not to do more than this.

The meaningfulness of the overall error-rate figures is, of course, questionable, for these figures do not take into account the nature of the errors made, nor the inaccuracies in measurement caused by cheating. Cheating undoubtedly made these figures an underestimation of the true error-rate. It was possible to limit it to a certain extent in the classroom by supervision, but it occurred fairly extensively in the homework sections towards the end of the programme (e.g., eighteen out of thirty-one pupils cheated on one problem in the penultimate lesson of the linear programme: this was detected by a misprint in the programme). Although the assumption cannot be held for every student that because homework was not supervised it was not conscientiously done, it may well be that this lack of control led to poorer test results obtained by the linear group than had been previously obtained during its development.

Such findings suggest that if difficulties in the programme are to be pinpointed, then control, either from a supervisor, or better, from a cheatproof device, is desirable. Without such control the measures of error-rate and time taken are distorted by cheating. This does not imply, however, that the results obtained are valueless. This experiment was conducted in (almost)

normal school conditions: it is in such conditions that programmes are most likely to be used in schools, and the results obtained are appropriate to them. What is implied, however, is that these conditions are limiting, and that the development of the original programme might have taken a different (and better) course if strict control over student responses had been achieved.

The results given in the section on student attitudes, indicating a significant decline in the popularity of each programme, are similar, but not as extreme as those found by Popham²⁰. The overall findings, after nine weeks, are similar, but not as clear cut as those of Beane² who compared the attitudes of pupils towards a linear and a multiple-choice branching programme on plane geometry. He found, as did the writer, that there was a slight preference for the linear programme, and that this preference was most marked in the high ability group.

Students may be very favourable to programmed instruction, but they may prefer other subject matters and approaches. This, in fact, was illustrated in this experiment. Initially student reactions to *Logarithms* were very favourable, but, (as responses to question six of the questionnaire indicated) this did not imply that conventional instruction in other subject matter was not enjoyed. Later, student reactions were much more neutral. In response to question six, however, *Logarithms* was ranked only one lesson lower: the programmed lesson was thus now considered much the same as any other. These results would seem to indicate the presence of a 'novelty factor' in the initial stages of the experiment: the effects of novelty, and its different forms, have been examined elsewhere²¹.

V.—CONCLUSIONS.

The results of this experiment are taken to support the contention that branching programmes—which are more sensitive to the needs of individuals—are superior to linear programmes. Such a conclusion is, of course, limited to the programmes and populations used. When the results are compared with those of other experiments in this field using skip programmes, it can be seen that a complete spectrum of possible results has now been achieved. The differences between the two programmes used in this experiment are so slight, however, that it may be legitimate to ask whether the experiment was worth doing. The answer to this question is positive, if only in a negative sense. We need experiments of this kind to indicate where progress may or may not be made, and to provide evidence to support our generalisations. The conclusions that the writer draws are that branching, where subjects are fairly sophisticated and are homogeneous in ability and preknowledge, is as effective as linear programming, and thus that comparison studies at this level may be a waste of time. However, the implication is that branching (particularly of the skip kind) where students are *not* homogeneous in sophistication, ability or preknowledge, may be more effective than linear programming. It is surprising that when one considers the experiments reported in this field to find that few studies in these conditions (i.e., of non-homogeneity) have been carried out.

ACKNOWLEDGEMENTS.—The writer would like to record his appreciation to Professor H. Kay and members of the D.S.I.R. unit on automated teaching in the Department of Psychology, the University of Sheffield, for constructive criticism of this paper; to D.S.I.R. for providing funds which allowed this research to be carried out; to the Sheffield Education Committee for permission to work in Sheffield schools; and to the Heads, teachers and pupils of schools in which this research was carried out.

VI.—REFERENCES

- ¹ COULSON, J. E., and SILBERMAN, H. F. (1960). The effects of three variables in a teaching machine. *J. Educ. Psychol.*, 51, 135-143.
- ² BEANE, D. G. (1962). A comparison of linear and branching techniques of programmed instruction in plane geometry. *Technical Report No. 1—Comparative Studies in Principles of Programming Mathematics in Automated Instruction*. University of Illinois, Urbana, Illinois.
- ³ ROE, A. (1962). A comparison of branching methods for programmed learning. *J. Educ. Res.*, 55, 407-416.
- ⁴ LARKIN, T. C., and LEITH, G. O. M. (1963). A comparison of linear and branching methods of programmed instruction. *Programming*, '63. Bull. Res. Unit. Dept. Education, University of Sheffield.
- ⁵ HERRINGSHAW, P. E., and HUNTER, J. H. (1964). Geography: branching *versus* linear programming. *Educ. Rev.*, 16, 109-119., University of Birmingham.
- ⁶ LARKIN, T. C. (1964). Electricity: a comparison between linear and branching methods of programming. *Educ. Rev.*, 16, 92-98. University of Birmingham.
- ⁷ GOLDSMITH, M. (Ed.) (1963). *Mechanisation in the Classroom*. London: Souvenir Press.
- ⁸ AUSTWICK, K. (Ed.) (1964). *Teaching Machines and Programming*. London: Pergamon.
- ⁹ CAMPBELL, V. N. (1962). Studies of by-passing as a way of adapting self-instructional programmes to individual differences. *Final Report A.I.R.—C41—5/62*. American Institute for Research, San Mateo, California.
- ¹⁰ GLASER, R., and REYNOLDS, J. H. (1962). Multi-tracking in programmed instruction: studies with a programme on mathematical bases for management decision making. *Programmed Learning Lab*. University of Pittsburgh, Pittsburgh, P.A.
- ¹¹ KNAPPER, C. K. (1964). Unpublished report. University of Sheffield.
- ¹² HARTLEY, J. (1964). *Logarithms*. London: Methuen.
- ¹³ GUILDFORD, J. P. (1956). *Fundamental Statistics in Psychology and Education*. New York: McGraw-Hill.
- ¹⁴ SIEGEL, S. (1956). *Non-Parametric Statistics*. New York: McGraw-Hill.
- ¹⁵ MARKLE, S. M. (1963). Faulty branching frames: a model for maximising feedback to the programmer. *Programmed Instruction*, III, 1, 4-5.
- ¹⁶ STOLUROW, L. M. (1961). *Teaching by Machine*. Washington, D.C.: U.S. Dept. Health Education and Welfare.
- ¹⁷ CAVANAGH, P. (1963). The Auto Tutor and Classroom Instruction: three comparative studies. *Occup. Psychol.*, 37, 44-84.
- ¹⁸ HARTLEY, J. (1964). *A Study in Programmed Learning*. Unpublished Ph.D. thesis, University of Sheffield.
- ¹⁹ BROOKS, L. O. (1961). *Response Latency in Programmed Learning: Latency Related to Error-Rate*. Unpublished Ph.D. thesis. University of Houston. Abstracted in *Audio Visual Communication Rev.*, II, 5 (1963).
- ²⁰ POPHAM, W. J. (1962). The influence of novelty effect upon teaching machine learning. San Francisco State College, San Francisco. Abstracted in SCHRAMM, W. (1964). *The Research on Programmed Instruction: An Annotated Bibliography*. Washington, D.C.: U.S. Dept. Health, Education and Welfare.
- ²¹ HARTLEY, J. (1965). Social Factors and Programmed Instruction. Paper accepted for *Programmed Learning*.

(Manuscript received 7th January, 1965.)

A STUDY OF PERSONALITY DIFFERENCES BETWEEN BLIND AND SIGHTED CHILDREN

BY HAMED A. S. ZAHKAN.

(University of London Institute of Education)

SUMMARY. In the literature there are two different points of view regarding the possible personality differences between the blind and the sighted. The first considers that blindness leads to compensatory behaviour, possibly accompanied by personality maladjustment and introversion. The second implies that in regard to basic personality variables the process of adjustment in blind persons is not significantly different from that of the sighted.

The writer carried out an experiment with matched groups of blind and sighted children to offer quantitative data on this problem. The Williams' Intelligence Test for Children with Defective Vision and Vernon's Abstraction Test of Intelligence were used for assessing the I.Qs. For personality assessment, the writer prepared a Blind Children's Structured Interview, a Sentences Completion Test and a Semantic Differential. The Junior Maudsley Personality Inventory was also used.

After a pilot study and administration of the tests, the data were analysed and almost all the test results agreed with each other in giving statistically non-significant differences in means in favour of the sighted, thus supporting the second point of view.

I.—INTRODUCTION.

PSYCHOLOGICAL guidance is particularly important for blind children. In the present study* one of the main objectives was to develop an understanding of psychological adjustment and counselling of blind children. The writer also attempted to develop personality measures which could be appropriate to testing blind children.

Examining the literature it was found that two different points of view are commonly held regarding the possible differences between the blind and the sighted. The first considers that a physical handicap, such as blindness, will lead to compensatory behaviour possibly accompanied by personality maladjustment or by introversion. This implies that the blind will exhibit a larger number of atypical responses to items of personality adjustment measures than a comparable group of sighted subjects. Examples of studies supporting this trend are by Brown (1938 and 1939), Dean (1958), Jervis (1959), and Dent (1962).

The second point of view suggests that, in regard to basic personality variables and dimensions, the process of adjustment in blind persons is not significantly different from that of sighted people. Examples of studies supporting this are by Sommers (1944), Scholl (1953) and Cowen and others (1961).

The writer offers some quantitative data on the problem which may help in the psychological guidance of blind children.

II.—THE POPULATION TESTED.

The population of the main study included 100 children in two groups : (1) A group of fifty blind children, (2) A group of fifty sighted children, who were matched for age, sex, I.Q. and socio-economic background according to the father's occupation.

* For the detailed study, see : ZAHKAN, HAMED A.S., 1964.

The age groups of the study were chosen to represent the upper classes of the primary stage and the lower forms of the secondary stage, i.e., from 9 years to 13 years 11 months. The original number of the subjects of the blind sample was eighty-seven, being the whole school population between these ages in an all-age residential school for the blind in Kent. In addition, thirteen children randomly chosen from the Sheffield School for the Blind made a group of 100 blind children. All of them were congenitally blind, educationally blind*, and institutionalized and had no additional or secondary handicaps. Such cases, in fact, represent the majority of blind children of school age.† It was intended that an equal number of sighted children would form the control group of the investigation, however, it was more convenient to keep the controls to fifty children drawn from a primary and a secondary modern school in Middlesex which could be considered one school since they were in the same building, and to select fifty from the blind group using a Table of random numbers. Nevertheless, the original group was found useful in supplying general data about education, attitudes, needs and problems in a structured interview designed for this purpose. The two groups consisted of equal number of both sexes.

III.—THE TESTS

The Williams' Intelligence Test for Children with Defective Vision, which is the most widely used test throughout this country, was used in testing the intelligence of the blind group (Williams 1956). It includes some items from the Terman-Merrill 1937 revision, chiefly from form M, a number from Valentine's Intelligence Test for children, some from Burt's Reasoning Tests, a few from group tests, modified slightly for individual testing, and the Vocabulary Tests from Wechsler's Children's Scale.

An Abstraction Test of Intelligence, constructed by P. E. Vernon was used in testing the sighted group. It includes a variety of items such as word, letter and number series. The subject is provided with a certain number of dots in each item and is instructed to write the missing letters or numbers wherever he sees a dot, e.g. (11 22 33 44 ..). Professor Vernon considers such a test to illustrate Spearman's education of relations and correlates. He also reports that its dependence on *g* is very high and at the same time its dependence on verbal ability or education is very low.

The writer prepared a "Blind Children's Structured Interview" aiming at obtaining first-hand data concerning each child's home and family, school, his needs, his interests, his general problems, his adjustment and his attitudes. The interview was conducted in the form of free discussion with the child. Examples of items are:

Do you think that you have more troubles than most blind boys and girls you know?

Do you think that you have more trouble than most sighted boys and girls you know?

Do you feel ashamed of your blindness?

Would you like some help with your personal troubles?

If you are troubled, to whom would you go for help? etc.

* The 1944 Education Act defines them as those who have no sight or whose sight is or is likely to become so defective that they require education by methods not involving the use of sight.

† Ministry of Health, 1963, Register of the Blind, National Statistics (England and Wales), London, H.M.S.O.

In view of the dependence of personality inventories on the subjects' attitudes (Vernon, 1964), the various projective approaches to personality were considered. Pictorial techniques would be inapplicable with blind children, hence it was decided to develop a Sentence Completion Test which could be given orally as a projective test to obtain an overall adjustment score and to locate specific adjustment difficulties within five broad areas of adjustment. Sentence stems were structured around adjustment in the following areas: Home adjustment, Social adjustment, Emotional adjustment, School adjustment, and Physical Health adjustment. Examples of the sentence stems are:

In scoring, all the responses were examined and classified into

- The writer also developed a Semantic Differential to give an objective measure of the connotative meaning of ten selected concepts as perceived by the blind and the sighted subjects. This method was developed by Osgood and others (1952, 1954, 1957 and 1962), for studying the connotative meanings that people apply to different concepts. A subject is given a list of concepts such as myself, my father, school, etc., and asked to rate each one on a series of adjectival scales, e.g. strong vs. weak, quick vs. slow, etc.

- (1) Evaluative,
- (2) Activity,
- (3) Potency.

Ten concepts were chosen to fall in the following areas:

Three Home concepts, (Home, My Father, and My Mother),

Two School concepts, (Teachers, and Lessons),

Two Social concepts, (Parties, and Other Sex), and

Three Self-concepts, (My Dreams, Myself Now, and Myself When I'm Grown Up).

Each item (i.e., pairing of a specific concept with a specific scale) presents the following situation:

[illegible]

- (1) Extremely X (e.g., Very Good).

- (2) Quite X (e.g., Quite Good).
- (3) Slightly X (e.g., Slightly Good).
- (4) Neither X nor Y, Equally X and Y (e.g., Neither Good nor Bad, Equally Good and Bad).
- (5) Slightly Y (e.g., Slightly Bad).
- (6) Quite Y (e.g., Quite Bad).
- (7) Extremely Y (e.g., Very Bad).

Thus, on these scales the subject's checks indicate both the direction and the intensity of his association.

In testing blind subjects, the writer devised a seven-step wooden scale fixed on a board. The same scale was used with the sighted subjects. It was also found useful to construct ten informative cards which could be fixed easily on the scales and the board. The positive and the negative adjectives of the scales were typed on these cards for the sighted and printed in Braille for the blind. This helped the subject to follow the change in positions of the positive and negative ends of the scales. The subject fixes his fingers on the scale using three fingers of his left hand and three fingers of his right hand. The middle space is left where thumb can be used.

Pilot Study: A brief pilot study preceded giving the tests which the writer had devised; this resulted in some rearrangement, rephrasing and modification of the tests.

Testing Procedure: In general the tests were given individually in a special room free from noise and interruptions, with no member of the school staff present. Only the Abstraction Test of Intelligence was administered as a group test since it was necessary to test all the children in both the primary and the secondary schools falling in the age range of 9 years to 14 years to choose fifty children among them having equivalent I.Q.s. to the blind group. It was difficult to give this test individually since the number of children tested was over 300.

IV.—ANALYSIS OF DATA

Data obtained from the tests were analysed. The Semantic Differential in particular was factor analysed.

On the Junior Maudsley Personality Inventory, the results show that many of the blind subjects tended to be more introverted than the sighted. The means being 11.52 and 12.30. This difference was not statistically significant (C.R. 1.59). There was no significant difference between boys and girls in either blind or sighted group. On neuroticism, a larger number of the blind subjects tended to score high. The means being 8.04 and 7.02. This difference in favour of the sighted again was not statistically significant (C.R. 1.59). Taking sex into account, it was found that blind boys as well as blind girls tended to score higher than the sighted. In both groups the girls scored lower on neuroticism than boys. The notion that females tend to score higher than males on neuroticism was not confirmed in this study. Comparing Furneaux's sample with our two groups we find that the means of our sighted group are almost the same as Furneaux's on both Extraversion and Neuroticism, whereas the blind group's mean on Extraversion was lower and their mean on Neuroticism was higher than Furneaux's sample.

TABLE 1

MEANS AND STANDARD DEVIATIONS OF FURNEAUX'S SAMPLE COMPARED WITH OUR TWO GROUPS.

	Furneaux's sample		Blind group		Sighted group	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Extraversion	12.39	3.46	11.52	2.53	12.30	2.39
Neuroticism	7.35	3.54	8.04	3.34	7.02	3.06

On the Sentence Completion Test, a larger number of the blind subjects had low scores—indicating that they were less adjusted than the sighted. The means being 6.88 and 8.54. This difference was not significant (C.R. 1.11). With regard to sex, there was no difference between blind and sighted girls. A clear difference was found between blind and sighted boys, the sighted scoring higher than the blind. It is also noticeable that in all the areas except Home adjustment, the blind group means are lower than those of the sighted, especially in the Emotional area.

The data taken from the Semantic Differential were given to the University of London Computer Unit to calculate intercorrelations and principal components:

- (a) for the blind children,
- (b) for the sighted controls.

Having ten variables (i.e., the ten scales), the score matrix for each analysis in fact was 500×10 .

A score matrix was drawn from each analysis = $c \times s \times n$ where c =concepts, s =scales, and n =subjects. Each subject provided a complete set of ten judgments on each of ten concepts, one judgment on each scale. Since both subjects and concepts are replicated it was possible to obtain separate matrices of scale intercorrelations for individual subjects (summing over concepts) as well as for individual concepts (summing over subjects). However, for the purpose of the present analysis of the Semantic Differential scores, it was then decided to sum over both subjects and concepts and to analyse the correlations between the ten scales.

Inspection of the correlational matrix leading to our factor analysis suggested the presence of a general factor which made all correlation coefficients positive.

Hotelling's principal component analysis method was applied to the matrix of correlations. The aim of obtaining principal components is to reduce the meaning of each concept, as conveyed by ten adjectival scales, to a few major dimensions which represent the average attitudes of the subjects.

Four principal components were extracted from the first analysis (Blind) and five principal components were extracted from the second analysis (Sighted). However, it was found difficult to interpret the component loadings as they stand without rotation, though certain groupings similar to those claimed by Osgood appeared to be present. Thus the extracted components were rotated into simple structure.

The results for the blind were found to yield different structure from those for the sighted. For the blind, there were three satisfactorily significant factors and the fourth factor did not contribute to the interpretation of the scales, so rotation was confined to the first three. For the sighted, all the five factors were rotated.

In the blind group, the first rotated factor was labelled the General Evaluative Factor. Its high loadings are on good-bad, happy-unhappy, nice-awful, successful-unsuccessful, interesting-dull, and wise-silly. The second factor is an Activity Factor, since the scales which had high loadings on it are big-little, quick-slow, and hot-cold. The third factor appears to be a Potency Factor. The most distinctively loaded scales are strong-weak, and big-little. Note however, that there is an overlap between the second and the third factors on big-little.

For the sighted group, the first factor was also found to be a General Evaluative one, by listing scales which had high loadings on it, good-bad, happy-unhappy, and nice-awful. A prominent group factor was found to be a Cleverness Factor, the scales which had high loadings on it being successful-unsuccessful, interesting-dull, and wise-silly. This factor could not be separated off in the blind analysis. The third factor was similar to that for the blind, namely a Potency group factor. Scales which had high loadings on it were strong-weak and big-little. The fourth and the fifth factors were found to be more specific. The fourth was an Activity Factor specific to quick-slow. The fifth factor represents another type of Activity being specific to hot-cold.

TABLE 2
THE ROTATED PRINCIPAL COMPONENTS LOADINGS (BLIND GROUP).

Scales	Loadings for 3 Components*			h ²
	I ²	II ¹	III ¹	
1. Good-bad727	.046	.231	.584
2. Strong-weak639	— .221	.614	.834
3. Happy-unhappy736	.227	.257	.660
4. Nice-awful837	.205	.015	.743
5. Big-little384	.524	.622	.810
6. Quick-slow656	.412	.136	.619
7. Hot-cold434	.730	.122	.736
8. Successful-unsuccessful	.794	.185	— .078	.671
9. Interesting-dull797	.262	.088	.712
10. Wise-silly828	.185	— .026	.721

* The number of superscripts shows the number of rotations of each component.

TABLE 3

THE ROTATED PRINCIPAL COMPONENTS LOADINGS (SIGHTED GROUP).

Scales	Loadings for 5 Components					h ²
	I ⁴	II ³	III ⁴	IV ⁴	V ³	
1. Good-bad823	.066	-.152	.209	.031	.749
2. Strong-weak471	-.004	.667	.274	-.148	.764
3. Happy-unhappy871	-.160	.047	-.168	.076	.821
4. Nice-awful824	.203	-.117	.135	-.016	.752
5. Big-little463	-.126	.732	.089	.206	.817
6. Quick-slow369	-.058	.195	.851	-.090	.909
7. Hot-cold418	.085	.161	.285	.817	.957
8. Successful-unsuccessful	.597	.560	.161	.177	.122	.743
9. Interesting-dull592	.590	.263	.067	.078	.779
10. Wise-silly542	.473	.128	.384	.138	.700

Thus, it is evident that the factorial structure in the two groups is different. The semantic space in the blind group is less structured than in the sighted. In order to get comparable factor scores in the two groups it was decided to take into account only the General Evaluative factor, scored on good-bad, happy-unhappy, nice-awful, quick-slow, hot-cold, successful-unsuccessful, interesting-dull, and wise-silly, and the Potency group factor scored on strong-weak, and big-little, since these factors were identified in both groups. This leads to simplifying the semantic space on to two dimensions.

The results of this factor analysis of the Semantic Differential suggest that the two groups differed in terms of the underlying dimensions of judgments they used in differentiating among concepts. They differed in the number of factors required to account for their judgments as well as in the relative weights referring to the same set of factors. It is also important to note that the two groups differed in the dominant Evaluative factor and that the blind subjects seemed to rely more heavily on general evaluation.

Calculating the mean factor scores for Evaluative and Potency factors, there was a clear similarity between the two groups in perceiving "Father, Home, Parties, Lessons, Dreams, Other Sex, and Teachers." The two groups differed slightly in perceiving "Mother." A marked difference was found in Self evaluation between the two groups "Myself Now and Myself When I'm Grown Up," which was statistically significant. (See Figures 1 and 2.)

Considering similarities and differences between the sexes, we find general similarity between blind boys and sighted boys except that the blind boys more strictly rejected blind girls and perceived them as definitely bad and weak, while they were more valuable and a little stronger for the sighted. Blind and sighted girls perceived 'Other Sex' more favourably than did the boys. It may seem that the Semantic Differential provides more suggestive information about the attitudes of blind and sighted children than did the single scores on accepted personality tests.

Almost all the test results agreed with each other in giving differences in means in favour of the sighted, the significance of these differences were not significant, apart from differences on the Semantic Differential indicative of Self-Evaluation and confidence in the future.

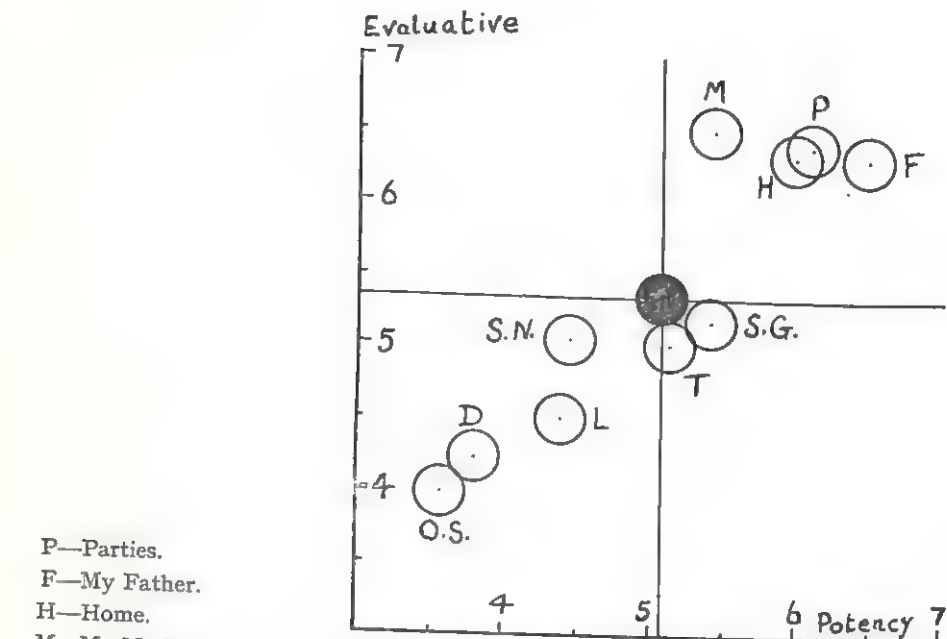


FIGURE 1
SEMANTIC SPACE FOR BLIND GROUP.

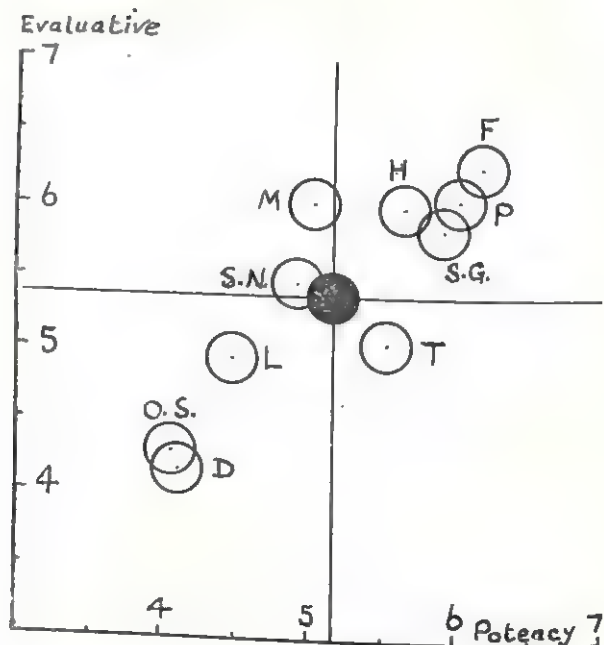


FIGURE 2
SEMANTIC SPACE FOR SIGHTED GROUP.

V.—DISCUSSION AND CONCLUSIONS.

Our experimental work was followed by a survey of what is being done for blind children in England and Wales with special attention to guidance services. As a step towards a proposed plan of a guidance programme and depending mainly upon the results of the Blind Children's Structured Interview and the writer's personal observations, some characteristics of blind children were analysed. These included basic needs, interests, problems, adjustment and attitudes.

Finally, the findings of the study were considered as a necessary background for the development of a flexible outline of a psychological guidance programme for schools for the blind.

The main conclusion of the study was that although blind children deviate in sensory characteristics, they possess much the same personality characteristics, the same drives, motives, needs and capacities as the sighted. There are no distinct personality problems produced by blindness, but problems frequently arise from the reactions of the blind to their social environment. This confirmed that blind children are and should be looked upon as ordinary children who cannot see. This also calls for the need for research in the field of personality and personal adjustment of blind children, particularly for studies of effective methods for changing public attitudes towards the blind and for experimentation in educating blind children in ordinary schools.

ACKNOWLEDGMENTS.—Hearty acknowledgments are due to Professor Philip E. Vernon who gave me regular guidance, direction, inspiration and encouragement.

I am very grateful to Mr. R. P. Bolton, Mr. F. H. G. Tooze, Mr. M. Price and Mr. A. B. H. Brigden, headmasters of the primary and secondary schools where the experimental study was conducted. Hearty acknowledgments are also due to their staffs.

The blind and sighted pupils who answered the tests deserve thanks for the wholehearted way in which they entered into it.

I am also grateful for the help I have received from the Library staff of the University of London, the Institute of Education, and the Royal National Institute for the Blind. My thanks are also due to the University of London Computer Unit staff for their help.

I should like to express my gratitude to Miss M. Williams, Mr. W. D. Furneaux, Mr. T. R. Holland, and Mrs. M. Morris as well as my wife.

I should like to express my sincere appreciation to the Government of the U.A.R. for the grant which has made possible this research.

VI.—REFERENCES

- BROWN, P. A. (1938). Responses of blind and seeing adolescents to an introversion-extraversion questionnaire. *J. Psychol.*, 6, 137-147.
- BROWN, P. A. (1939). Responses of blind and seeing adolescents to a neurotic inventory. *J. Psychol.*, 7, 211-221.
- COWEN, E. L., UNDERBERGE, R. P., VERILLO, R. T., and BENHAM, F. G. (1961). *Adjustment to Visual Disability in Adolescents*. New York: American Foundation for the Blind.
- DEAN, S. I. (1958). Some experimental findings about blind adjustment. *New Outlook for the Blind*, 52, 182-184.
- DENT, O. B. (1962). An investigation of attitudes toward work adjustment of the blind. *New Outlook for the Blind*, 56, 357-362.
- EYSENCK, H. J. (1947). *Dimensions of Personality*. London: Kegan Paul.

- EYSENCK, H. J. (1952). *The Scientific Study of Personality*. London: Routledge and Kegan Paul.
- EYSENCK, H. J. (1957). *The Dynamics of Anxiety and Hysteria: An Experimental Application of Modern Learning Theory to Personality*. London: Routledge and Kegan Paul.
- EYSENCK, H. J. (1959). *Manual of the Maudsley Personality Inventory*. London: Univ. of London Press.
- EYSENCK, H. J. (Ed.) (1960). *Experiments in Personality*. Vol. I. *Psychogenetics and Psychopharmacology*, and Vol. II, *Psychodiagnostics and Psychodynamics*. London: Routledge and Kegan Paul.
- EYSENCK, H. J. (1962). *Uses and Abuses of Psychology*. London: Pelican.
- EYSENCK, H. J. (1963). Comment on the relation of neuroticism and extraversion to intelligence and attainment. *Brit. J. Educ. Psychol.*, 33, 192.
- FURNEAUX, W. D., and GIBSON, H. B. (1961). A children's inventory designed to measure neuroticism and extraversion. *Brit. J. Educ. Psychol.*, 31, 204-207.
- JERVIS, F. M. (1959). *A Comparison of Self Concepts of Blind and Sighted Children, Guidance Programmes for Blind Children. A Report of a Conference*. Edited by Davis, C. J., Watertown, Mass., Perkins School for the Blind, 19-25.
- JUNG, C. G. (1923). *Psychological Types*. Translated by Baynes, H. G. Harcourt: Brace and Co.
- OSGOOD, C. E. (1952). The nature and measurement of meaning. *Psychol. Bull.*, 49, 197-237.
- OSGOOD, C. E. (1962). Studies on the generality of affective meaning systems. *Amer. Psychologist*, 17, 10-28.
- OSGOOD, C. E., and LURIA, Z. (1954). A blind analysis of a case of multiple personality using the semantic differential. *J. Abnorm. Soc. Psychol.*, 49, 579-591.
- OSGOOD, C. E., SUCI, G. J., and TANNENBAUM, P. H. (1957). *The Measurement of Meaning*. Urbana, Ill.: Univ. of Illinois Press.
- SCHOLL, G. (1953). Some notes on the use of two personality tests with visually hand capped students. *New Outlook for the Blind*, 47, 287-295.
- SOMMERS, V. S. (1944). *The Influence of Parental Attitudes and Social Environment on the Personality Development of the Adolescent Blind*. New York: American Foundation for the Blind.
- VERNON, P. E. (1963). *Personality Assessment: A Critical Survey*. London: Methuen.
- VERNON, P. E., and PARRY, J. B. (1949). *Personel Selection in the British Forces*. University of London Press.
- WILLIAMS, M. (1956). *Williams' Intelligence Test for Children with Defective Vision*. Birmingham: Univ. of Birmingham, Institute of Education.
- ZAHRAH, H. A. S. (1964). *An Investigation of Some Aspects of Personality with Reference to the Psychological Guidance of Blind Children*. Unpublished M.A. thesis, University of London.

(Manuscript received 7th January, 1965.)

THE EFFECT OF LOGIC INSTRUCTION ON THE VALENTINE REASONING TEST

By CHARLES F. ELTON
(*University of Kentucky, USA*)

SUMMARY. Students taking a sophomore course in logic and students taking a sophomore course in applied psychology were administered the Valentine Reasoning Test during the first and last weeks of the semester. A test was made of the null hypothesis that there would be no difference in the test score due to type of instruction. The null hypothesis could not be rejected. The implication for the use of this test in the admission of advanced transfer students is indicated.

I.—INTRODUCTION.

THE evolution and standardization of a new "Reasoning Test for Higher Levels of Intelligence" have been described by Professor C. W. Valentine (1961). In the development of his test, Professor Valentine (1961, p. 228) was guided by the factor analysis research which indicates that the ability to reason is dependent upon "g." The sixteen items in his test are divided into four inductive reasoning problems and twelve deductive reasoning problems. Hallworth (1963), using Professor Valentine's data, discovered by the method of principal components a general reasoning factor. However, this general factor accounted for only 22 per cent. of the variance for males and 18 per cent. for females (Hallworth, 1963). A Varimax rotation of the axes produced the additional finding that three types of reasoning existed among the items: intuitive, inductive, and complicated (Hallworth, 1963). Hallworth also reports that these factors were not uniform for the two sexes, but further support for the three factors was obtained by a Quartimax rotation.

Professor Valentine (1961, p. 231) argues cogently that his test may be especially useful in university admission, e.g., "In view of its relative independence of educational influences, it would reveal the highly intelligent individual who had had unsatisfactory schooling, and on the other hand, the dull student who had been excessively coached and crammed." Professor Valentine (1961, p. 231) also states: "I think it probable that a training in symbolic logic or in the detection of syllogistic fallacies might have an appreciable effect on the scores in Section B: but such a training is likely to be exceedingly rare among boys and girls at school competing for admission to training colleges or universities." It is very clear from the context that Professor Valentine is referring to the initial admission of a student to a college or university. While it may be reasonable to expect that students will not have had training in symbolic logic prior to university matriculation, is this to be expected for university students who, after their initial entrance, transfer from one institution to another? Obviously not, and should training in logic influence test results there would exist a serious limitation to the use of the Reasoning Test as a selection instrument for advanced transfer students.

It is the purpose of this study to assess Professor Valentine's assumption that training in logic might influence the score a student obtains on the Reasoning Test.

II.—PROCEDURE.

During the first week of class in the fall semester of 1963, all students enrolled in Philosophy 220—Elementary Logic, a beginning course devoted to the logical principles underlying valid thought, were given the Valentine Reasoning Test according to the handbook of instructions (Valentine, 1954). A total of sixty students were enrolled in three sections of this course, taught by two professors. The same textbook was used in each section, namely, *Introduction to Logic* by Irving M. Copi, of the University of Michigan.

During the same week forty-six students enrolled in a sophomore course in applied psychology were given the test. Both of these courses are, according to University of Kentucky academic regulations, closed to freshmen students. The professors volunteered the use of the students in their classes without knowing the purpose of the experiment. The experimenter told the students he was interested in developing test norms for American students, the results would be confidential, and the scores would not and could not be used in assigning course grades.

During the last week of classes the test was again administered to these same student groups. However, by this time, due to course cancellations, only forty-six students were enrolled in the logic course and forty-six students were still enrolled in the applied psychology course. Only one student in the control group had completed the beginning logic course. The resulting data were subjected to a three factor analysis of variance done on the IBM 7040 computer.*

* The author wishes to express his gratitude to Mr. Selwyn Zerof of the University of Kentucky Computing Center for the statistical analysis.

III.—RESULTS AND DISCUSSION.

The results are given in Table 1. It is clear from Table 1 that the large error variance between subjects taking logic and those taking applied psychology precludes a significant difference. Hence, the assumption that logic instruction affects the scores on the Reasoning Test is not tenable. This finding, of course, must be qualified in that it is true only for a heterogeneous sample of non-freshmen undergraduate college students who were taught over a period of sixteen weeks by only two different Assistant Professors in an American land-grant university.

TABLE 1
ANALYSIS OF VARIANCE OF REASONING TEST SCORES.

Source of Variation	DF	MS	F
Between Subjects :	91		
Groups (A)	1	190.83	
Subjects w groups	90	165.53	
Within Subjects :	276		
Test-retest (B)	1	362.02	10.05*
AB	1	10.11	
B x subjects w groups	90	36.00	
Part of test (C)	1	6089.70	87.07*
AC	1	.98	
C x subjects w groups	90	69.94	
BC	1	111.98	
ABC	1	66.98	
BC x subjects w groups	90	31.02	

* —.01 level.

The significant main effect due to the test-retest condition (B) is a likely function of practice effect. The novel test items aroused unusual interest and it is known that some items, especially the first four, were remembered and discussed among the students as they left the testing situation. Since the AB interaction effect was insignificant it is reasonable to believe that the practice effect operated similarly among the students taking logic and psychology. This result, however, does emphasize the need for an alternate form of the test. The significance of the main effect due to the two parts of the test (C) is not unexpected since it is possible to obtain a maximum score of 23 points on the first part of the test and 48 points on the second part of the test. The lack of significance of the AC interaction effect indicates that neither the logic nor the psychology students differed markedly in the number of points earned on each part of the test. The fact that the BC interaction effect approaches the .01 level of significance supports the notion that there is apt to be little uniformity in test-retest score between the two parts of the test. Again, it is hypothesized that this type of interaction effect would not have occurred had an equivalent form of the test been administered in the retest condition.

IV.—CONCLUSION.

The selection of university students is an urgent and pervasive issue in higher education. The sheer magnitude of numbers has directed most research and discussion toward beginning university students, and the selection of advanced students transferring between universities has received little attention. There are many students in this group whose college work is either from relatively unknown or new institutions and the quality of their work is neither clearly pass nor clearly fail. Thus there is an apparent need for an objective assessment of ability to learn, independent of previous educational influences. Test results either are not available to most university admission offices in these cases or, if they are available, they are inappropriate because they are based upon norms for beginning university students. It seems advantageous to be able to use the same admission test for beginning and advanced students. The Valentine Reasoning Test appears to be a most suitable selection instrument for both groups because of its demonstrated validity. That it is not greatly influenced by a beginning course in logic is an impressive argument for its use with advanced students transferring between universities.

V.—REFERENCES.

- VALENTINE, C. W. (1954). *Reasoning Tests for Higher Levels of Intelligence*. Edinburgh : Oliver and Boyd.
VALENTINE, C. W. (1961). The use of a new reasoning test for selection of university and training college students. *Brit. J. Educ. Psychol.*, 31, 227-231.
HALLWORTH, H. J. (1963). An analysis of C. W. Valentine's Reasoning Test for higher levels of intelligence. *Brit. J. Educ. Psychol.*, 33, 41-46.

(Manuscript received 11th January, 1965.)

A FOLLOW-UP OF TEACHERS FIVE YEARS AFTER COMPLETING THEIR TRAINING

BY S. WISEMAN AND K. B. START

(*University of Manchester School of Education*)

SUMMARY. Biographical questionnaires were completed by 248 teachers who had qualified at seven colleges and one university department within one a.t.o. in 1955. The subjects' head teachers provided confidential references which were treated by (a) content analysis, (b) independent rating by four judges. Quantified data from the references and from the questionnaires were inter-correlated and factor analysed. Separate analyses were made for teachers in different types of school. The first eight factors were subject to a varimax rotation. Four common factors emerged: headmaster's opinion of the teacher, attainment on training course, attendance at courses, and satisfaction with teaching. Little correlation was found between college assessment and the various criteria of success in the profession.

I.—INTRODUCTION.

IN 1960, the University of Manchester School of Education instituted a survey of all the teachers who qualified in its constituent institutions in 1955 and who by then could have had five years' teaching experience. The training course results for all these teachers were available from university records. The group provide a population of 694* who were circulated with a biographical questionnaire which included a request for the teacher's own estimate of his satisfaction with teaching as a profession and for causes of dissatisfaction. The teacher's permission was also sought to approach his headmaster or headmistress for a confidential reference. For those who were no longer teaching the questionnaire contained a section relating to reasons for leaving the profession and the dates of leaving and of possible return to teaching. Over a period of twelve months replies from 525 teachers were received, of whom 102 declined to permit the investigators to approach their headteacher for a confidential report and thirty-four had taught for only a few months before leaving teaching four years previously. Where such permission was granted the head teacher was invited to supply a confidential reference for that teacher, of the kind he would write were the teacher applying for another post. The head was also asked to rate the teacher on a seven-point scale. 113 head teachers did not reply, thus leaving a group of 276 teachers on whom reference data were available. On final checking it was found that twenty-eight subjects had not provided all data on their questionnaires and that complete data on all aspects of the questionnaire, background and headmaster's reference were available for 248 teachers.

II.—QUANTIFICATION OF DATA.

Four independent judges gave a general impression rating of the quality of the teacher from the reference provided by the head teacher. Rater inter-correlation was between .62 and .79 (mean .72). The judges were asked to repeat their ratings three months later. Individual reliabilities ranged from .70 to .89, and the reliability of the *aggregate rating* (used as a criterion measure) was .92.

* Two-year trained at colleges and one-year trained at university.

A content analysis of the head teachers' references was performed by one of the authors. A list was made of all favourable (positive) and unfavourable (negative) comments. It was found that these comments (over 1,000 in number) would be classified under seven heads: personality, teaching ability, cognitive ability, organizing ability, extra-mural activity, relations with head and staff, and relations with children and parents. Each reference was now re-scrutinized by one of the authors and a rating given (on a five-point scale) for each of these seven qualities. In order to reduce halo, all references were rated on *personality* before going on to deal with *teaching ability*, followed by *cognitive ability*, etc. The scores on these qualities were also summed to give a second over-all assessment of the teacher.

The biographical questionnaire was used to provide a measure of success in the profession which was independent of that derived from the aggregate rating of the head's reference, by estimating the *degree of responsibility of present post*. This ranged along the continuum assistant teacher, extra responsibility (unpaid), graded post, deputy head, head.

The other information on the questionnaire (see Appendix) was also quantified. The qualifications of the teacher when he entered his training course was graded from the minimum of five 'O' level G.C.E. passes through 'A' level passes to (in the case of the graduates in the university department of education) the degree class. Also noted were the qualifications of the college-trained teacher when he applied for entry, as distinct from when he actually entered college.

Membership of organizations pertaining to education were assessed into 'professional' and 'academic' groups, as might be exemplified by the N.U.T. and the Science Masters' Association respectively. Within each category the participation of the teacher was assessed on the basis of any position held in the group.

Any courses followed by the teacher, subsequent to his qualifying in 1955, were classified into those leading to further qualifications and those which did not, e.g., Dip.Ed. or a six-lecture course on the new teaching alphabet, respectively. The non-qualification courses themselves were sub-divided as professional or personal. Thus, a course in backward children would be considered professional, whereas one in flower arrangement would be considered personal. Naturally this assessment was not clear cut, as what might be a professional course for one teacher might be a personal one for another. Where any doubt existed, the criterion adopted was whether the course appeared to be associated with the teacher's responsibilities in school. If so it was considered professional but when such a relation could not be seen it was listed as personal. Each course was also assessed for its duration and the total time spent attending courses was graded on a 5-point scale.

The teachers had been asked to grade the degree of their satisfaction with their job on a five point scale and this, together with the number of items the teacher listed as causing dissatisfaction, were taken as an estimate of personal contentment in the profession. The intensity of any one issue causing dissatisfaction was not assessed.*

The performance of the teacher during his training was based on his scores on the final theoretical papers in Principles of Education and the average he obtained in all his theory papers. The final teaching mark was also noted.

* For a fuller analysis of this variable, see: RUDD, W. G. A., and WISEMAN, S. Sources of dissatisfaction among a group of teachers. *This Journal*, 1962, 32, 275-291.

The student's age on entry to his training course was graded on a five-point scale and whether or not the teacher was still teaching at the time of the survey was noted.

III.—TREATMENT OF DATA.

The returned questionnaires were not from the complete population of the course studied (1953-55). Neither was it an organized sample based on statistical procedure. Rather it was determined by the ability of the investigators to contact all the group and the willingness of the members of the group to assist in the survey. All inferences drawn from the data must be made with this in mind. The first examination of the data was to discover if the teachers who did not reply differed in their entrance and training scores from those who did. Significant differences were not found on any of the data available. Secondly, the difference between the teachers who permitted an approach to be made to their headmaster and those who did not also failed to reach statistical significance. Thus, from the data available the three groups could have come from the same population and differences resulting in the variation in their responses to the survey might be attributable to unmeasured traits but not by the data considered in this study.

The data were grouped in a variety of ways. Initially teachers were grouped according to whether their present post was in a primary, secondary modern or grammar school. Two other groupings were considered—graduates, and two-year trained teachers who were on the staff of secondary modern schools. These five groupings of data were analysed on the Atlas Computer of the University of Manchester. They were inter-correlated, and then subjected to a principle components analysis. The first eight resulting factors were given a varimax rotation to simple orthogonal structure. The factor structures of the sub-groups were compared and four factors common to all types of schools and to the graduate and non-graduate secondary modern teacher appeared. Other factors appeared to be common to various combinations of some of the five groups but not obviously in them all.

IV.—COMMON FACTORS.

A Common Factor, loading heavily on the assessments of the head teacher's report, was obtained as the first factor in four of the five groups, and for the fifth, the graduates, it was the second factor abstracted. The items on which this factor loaded are given in Table 1 in which only those variables with at least one loading greater than .2 are included. (Complete loadings on all factors are given in the Appendix.)

This factor could be called *Head Teacher's Estimate of the Teacher*. It is notable that college assessments have little relation to this estimate of their product after five years' practising the profession for which the college prepared them. This is true of both the theory assessments of the teacher and the rating of his performance on teaching practice. For example, the first order correlations of the subjects' teaching practice mark with the *teaching ability* estimate for grammar, secondary modern and primary teachers respectively were .195, .175, .138, with *responsibility of present post* were .338, .127, .105, with the *aggregate rating* of the four judges were .335, .303, .077, and with the *head master's rating* were .280, .283 and .071

TABLE 1
LOADINGS ON COMMON FACTOR 1.

Variables	Graduate	Gram-mar	All Sec. Mod.	Non-grad. Sec. Mod.	Primary
1. Aggregate Rating of reference ..	.758	.695	.805	.782	.853
2. Personality796	.773	.769	.753	.850
3. Cognitive Ability588	.298	.353	.216	.348
4. Organizing Ability320	.264	.628	.389	.484
5. Teaching Ability840	.687	.771	.777	.845
6. Relations with Head and Staff ..	.725	.882	.678	.627	.729
7. Relations with Parents and Children613	.839	.595	.626	.698
8. Extra-Mural Activities399	.173	.675	.406	.360
9. Total 2-8962	.907	.980	.885	.959
10. Head Teacher's Rating485	.451	.693	.626	.646
11. Responsibility of Present Post ..	.038	.012	.219	.039	.102
12. First School265	.236	-.055	-.048	-.071
13. Number of previous posts	-.041	-.242	-.131	-.150	-.129
14. Teaching/non-Teaching	-.249	-.281	.361	.056	.020
15. No. Personal Courses	-.001	-.064	.140	.331	.037
16. Professional Organizations123	.085	.230	.126	-.127
Common Variance	15.65	16.25	19.29	16.33	18.29
Factor Order	2	1	1	1	1

Perhaps the most interesting aspect of this factor lies in the variations between school groups as to what aspects of the teacher the head master comments on. *Cognitive ability* seems valued among graduates but is not so highly regarded among the other groups. The difference between graduate and other groups, particularly grammar, is interesting in the light of the relative distributions of graduate and non-graduates in the grammar and secondary modern school. The *organizing ability* of the staff is frequently mentioned and valued highly by the head of the secondary modern school and, to a lesser extent, by his primary school colleague, but such an interest was not in evidence among head masters of the grammar or graduate teachers. Similarly, *extra-mural activity* is valued highly in the secondary modern school and little in the grammar school. The discrepancy between grammar and graduate groups here may be explained by the stressing of the extra-mural activity of the non-graduate teacher in the grammar school when the head master seems at a loss to say anything else about him.

Common factors 2 and 3 varied between the different groups in their rank order of variance size. One factor, with a mean variance of 11.33, appeared once as the first, once as the second and three times as the third factor abstracted. The other factor has an average variance of 10.29 and appeared three times as the second and twice as the third factor abstracted. The second common factor was taken to be the one with the highest average variance. This had high loadings on attainment during the training course and might be considered as the *training attainment of the teacher*. The item loadings are given on Table 2, listing all variables loading .2 or above on any one factor.

TABLE 2
COMMON FACTOR 2—"COLLEGE ATTAINMENT."

Variables	Graduate	Gram-mar	All Sec. Mod.	Non-grad. Sec. Mod.	Primary
Principles of Education372	.814	.796	.929
Teaching Practice501	.829	.725	.783	.736
Average Practice921	.473	.817	.796	.902
Average Theory862				
Average Essay713				
Average + Practice932				
Age	-.004	.151	-.152	-.015	.254
Entry Qualifications274	.047	.125	.203
Application Qualifications299
Degree Class755				
Faculty639				
Responsibility of Present Post284	.449	.024	-.091	.059
First School065	.539	.131	.126	.046
Number of Previous Posts	-.067	-.174	-.064	-.071	-.250
Courses leading to Qualifications484	-.074	.157	.084	.158
Length of courses340	-.150	.178	.140	.036
Professional Courses279	.048	.216	.125	.226
Professional Organizations227	.119	-.025	.041	.068
Academic Organizations568	.133	.121	.137	-.182
No. Causes of Dissatisfaction307	.101	.117	-.082	.041
Four Judges' Rating	-.061	.391	.163	.182	-.086
Cognitive Ability402	.708	.333	.116	.144
Teaching Ability128	.457	.279	.114	.065
Relations with Head and Staff	-.123	-.190	-.322	-.489	.142
Extra-mural Activities	-.112	.075	-.074	-.461	.041
Total	-.047	.237	-.020	-.250	.100
Head's Rating	-.015	.448	.246	.221	-.120
Common Variance	17.61	10.64	8.86	9.85	9.68
Factor Order	1	3	3	3	2

Among future secondary modern teachers the lack of relation between the academic status of the student on entry to his training course and his performance during training is interesting in the light of the loadings for future grammar and primary school teachers. Also notable is the very high loading of *degree class* in the graduate training courses. The high loading of .639 for *Faculty* might suggest that either the Arts graduates find the teacher training courses more allied to their background than do the scientists, or that this reflects a tendency for better Arts than Science graduates to enter teaching. This remains to be determined. Finally, for predicting college performance of future primary teachers, the greater significance of the teacher's qualifications on *application* for admission to a teacher training course rather than on his *entry* deserves further investigation.

It is not unexpected to see that the better students had the better *present posts* in grammar schools and among graduates. However, attainment during the training course seemed to bear little relation to *responsibility in present post* for secondary modern and primary teachers. The difference between the grammar and secondary modern teachers is interesting. The absence of such a trend in the primary schools might be explained by the fact that the five-year period which had elapsed since completing training is relatively short for

promotion in the primary schools where the number of special responsibility allowances is much smaller than in secondary schools of all types. Among the present grammar school teachers those with highest training attainment were likely to be appointed to the more academic secondary schools on their first appointment. Relation between training, attainment and first school was apparently absent in the other groups. Very few teachers who at the time of the survey were in primary schools had been to secondary or other schools, so that this item showed little variance. The tendency to change posts weighed slightly negatively on all groups though most significantly in the primary school.

Attendance at courses leading to further qualifications seemed to be linked with good training performance among the graduates, but not among the other groups. On the other hand, attendance at *professional courses* (not leading to qualifications) seemed linked with training performance in graduate, secondary modern and primary groups.

In their references the head teachers' comments on the *cognitive ability* of the teacher are less linked with training performance for the primary and secondary modern than for the grammar school teacher. The fact that the loading for the grammar school teacher was higher than for the graduate teacher might be explained by the lack of such comment on the non-graduates in the grammar school. The higher loadings in the grammar than the graduate group is consistent in the headmasters' reports. Similarly, for the grammar school group the teacher's *teaching ability* as commented upon by his headmaster and also as estimated by his college five years earlier, loaded highly, being .457 and .829, respectively.

The relations between the secondary teacher and his head and his staff colleagues appear to be slightly below average if his training attainment is above average, with the trend being most emphatic among the college trained secondary modern teachers. Presumably this is accentuated by the decrease in academic orientation of the secondary modern school. Intriguingly, in the primary school a small positive relation is found.

Common Factor 3 loads heavily on items relating to *courses attended* by the teacher after qualifying in 1955. The item loadings on the different groups are given in Table 3.

There seems to be a general tendency in all secondary schools for the teachers in the more responsible posts to be the ones who attended most courses but this trend did not seem to exist in the primary school where attendance was unrelated to the responsibility of teachers' *present school position*. Among the secondary modern teachers age, *membership of academic organizations* and *number of causes of dissatisfaction* seem to be linked with attendance at post-training courses. This feature of the secondary modern teacher's pursuit of better qualifications might appear to lead to some friction with their heads as the headmasters' reports suggest that *relations with the head and staff* are less smooth (loading $-.149$ and $-.185$ for all secondary modern and college trained secondary modern teachers, respectively) and that the head teachers considered such staff to be lacking in the *personality attributes* hoped for in a teacher (loading $-.274$ and $-.427$, respectively, for the groups mentioned earlier). It is conjectural whether this is a true reflection of the teacher, is the result of the thwarted demands of the head teacher on his member of staff who had academic commitments elsewhere, or—perhaps more likely—represents the efforts of the dissatisfied teacher, at odds with his headmaster, to seek satisfaction (and promotion) along such channels.

TABLE 3

COMMON FACTOR 3—"ATTENDANCE AT COURSES."

Variables	Graduate	Gram-mar	All Sec. Mod.	Non-grad. Sec. Mod.	Primary
Courses leading to Qualifications ..	·683	·866	·817	·753	·866
Length of Courses	·848	·909	·841	·856	·858
Professional Courses	·620	·785	·510	·574	·535
No. Personal Courses	·661	·340	·358	·516	·234
Age	—·181	—·138	·555	·276	—·003
Entry Qualifications		—·451	—·137	—·113	·160
Principles of Education		·413	·177	·236	·023
Practice of Education	·140	—·079	·154	·220	·187
Average Practice	·217	·280	·170	·102	·082
Average Theory	·256				
Responsibility of Present Post	·324	·323	·281	·441	—·067
First School	—·184	—·339	—·183	—·128	—·015
No. of Previous Posts	·042	·228	—·025	·189	·015
Teaching/non-Teaching	·114	·160	·370	·222	·183
Professional Organizations	·037	—·018	·150	—·049	·354
Academic Organizations	—·168	·144	·492	·574	·148
No. of Causes of Dissatisfaction	—·045	·099	·292	·420	·134
Four Judges' Rating	—·212	—·153	·154	·228	—·018
Personality	·045	·005	—·274	—·427	—·064
Organizing Ability	·015	—·081	·148	·125	·263
Extra-Mural Activities	·411	·087	·184	·194	·117
Head's Rating	—·062	—·170	·170	·261	—·061
Common Variance	8·68	11·46	10·86	12·37	8·07
Factor Order	3	2	2	2	3

The contrast in loadings for grammar schools and primary schools on *entry qualifications* is interesting. It suggests a tendency for the *less* well qualified grammar school teacher and the *better* qualified primary teacher to engage in outside courses, and particularly professional courses. Perhaps the non-graduate staff feel some pressure to improve their qualifications and status within the grammar school situation. This view might be supported by the *Principles of Education* loading of ·413 which suggested that college trained teachers who did well theoretically and found themselves in grammar schools had the incentive to continue their search for qualifications, especially if they had moved from a secondary modern to a grammar school during their teaching career (loading on *first school*—·339).

Membership of professional organizations appeared an important item in this factor for the primary school teacher only. Similarly, *extra-mural activities* loaded highly only for the graduates. On this latter point it might be that the knowledge of the graduate teacher's furthering his training prompted his head teacher to mention it more as a significant feature of the teacher's extra-mural behaviour, and it did not appear to affect relations with the head or to cast doubt on his personality as in the case of the secondary modern teacher.

TABLE 4
COMMON FACTOR 4—"SATISFACTION."

Variables	Graduate	Gram- mar	All Sec. Mod.	Non- grad. Sec. Mod.	Primary
Degree of Satisfaction796	.863	.798	.811	.803
No. Causes of Dissatisfaction	-.449	-.742	-.709	-.655	-.773
Age435	.203	.345	.455	.133
Entry Qualifications		-.029	-.101	-.310	-.208
Application Qualifications					-.213
Principles of Education		-.113	-.221	-.192	-.076
Practice of Education422	.107	.118	.120	-.030
Average Theory	-.249				
Average Essay	-.218				
Responsibility of Present Post214	-.055	.296	.490	-.042
No. of Previous Posts	-.247	-.238	.016	.145	-.033
Professional Courses180	-.242	-.228	.066	-.066
Personal courses	-.052	-.033	-.183	-.236	-.203
Professional Organizations	-.006	-.198	-.107	-.104	.203
Academic Organizations	-.287	-.210	-.063	-.210	-.040
Four Judges' Rating264	-.082	.276	.281	-.021
Cognitive Ability042	-.093	-.164	-.083	-.325
Teaching Ability	-.065	-.092	.150	.205	.102
Head's Rating538	.194	.300	.358	-.111
Common Variance	6.58	6.32	6.46	8.01	5.96
Factor Order	4	6	4	4	6

Common Factor 4 loaded highly on items relating to *satisfaction*. The relevant data are given in Table 4. It seems that it is the older secondary school teacher who is more satisfied with the profession, whilst in the primary school there is a slight tendency for the older teacher to be less satisfied than his younger colleagues. There is a tendency among the graduates for teaching practice assessment to be linked with subsequent satisfaction, and yet all the theory items seem to have small negative loadings with satisfaction.

Satisfaction seemed linked with *present post* for secondary modern teachers but not for grammar or primary. The low positive loading among graduates probably reflects the contentment of those graduates in secondary modern schools with posts of responsibility bringing financial and other rewards. Movement from post to post among the graduates and grammar school teachers does not seem to favour satisfaction among the teachers, and *attendance at courses* not leading to a qualification is perhaps an outlet to relieve dissatisfaction among secondary modern and primary school teachers. The satisfied secondary school teacher appears to be the better teacher as assessed by the headmaster's rating and the rating of the individual from the head's report by four judges, though the loadings of the latter item are much lower than the head's direct rating. For the group of primary school teachers, this relation is insignificant and negative ($-.111$ with head's rating).

One of the most significant differences among school types is shown by *cognitive ability*. The loading of $-.325$ for primary school teachers suggests a disturbing dissatisfaction felt by the more intellectual members of primary school staffs.

After the extraction of four factors which could be considered common to all groups of teachers, the differences between groups in their factor structure made it difficult to see further significant common factor patterns. Obviously individual factors relating to teaching experience existed but there seemed no overall pattern which fitted every group and factors loading highly on the teacher's professional experience varied considerably from group to group.

V.—PRIMARY SCHOOL FACTORS.

In the primary school group three factors were extracted in the order 4, 5 and 7. Each of these loaded highly on one aspect of the teacher's experience to the virtual exclusion of the others: i.e., Primary Factor (P.F.) 4 loaded .676 on first school, P.F.5 .793 on number of previous posts and P.F.7 loaded .776 on present post. The loadings are summarized in Table 5.

TABLE 5

SUMMARY OF LOADING AMONG PRIMARY TEACHERS ON FACTORS 4, 5 AND 7, RELATING TO TEACHING EXPERIENCE.

Variables	P.F.4	P.F.5	P.F.7
Responsibility of Present Post217	.014	.776
First School676	.050	.092
No. of Previous Posts	-.031	.793	.213
Age	-.342	-.369	.577
Application Qualifications655	.401	-.230
Entry Qualifications605	.371	-.367
Teaching/non-Teaching	-.069	-.476	.198
Professional courses076	-.058	.483
Personal Courses207	-.010	-.146
Professional Organizations326	.139	.063
Academic Organizations482	-.263	-.259
No. Causes of Dissatisfaction113	-.246	.045
Cognitive Ability	-.111	.321	.084
Organizing Ability	-.260	-.010	-.142
Relationship with Children/Parents	-.200	-.124	.289
Extra-Mural Activities266	-.469	.215
Head's Rating126	-.265	-.243
Teaching Ability037	.014	.090
Total % Variance	7.50	7.07	5.82

The most important of these is P.F.7, which shows the variables which are important to the primary school teacher in achieving promotion within the profession. *Age* is the most important, although in an area training organization containing one day college* specially designed for the training of mature women the possibility that this is a concealed *college* factor cannot be ignored. The *number of professional courses* is an important element in promotion, although it is negatively related to *qualifications on entry*. As might be expected for primary school teachers, *relationships with children and parents* is important for advancement, as is *extra-mural activities*. It is odd, however, that this should be important than *teaching ability* (loading .090) and *cognitive ability* (.084)—and even odder that promotion is negatively correlated with the head teacher's rating (−.243).

* At the period the subjects were in training. Since then other day training colleges have been opened.

VI.—SECONDARY SCHOOL FACTORS.

Among the college trained secondary modern teachers' experience seems to load highly on two factors, one relating to first school ($\cdot788$) which also has a moderate negative loading on *number of previous posts*, and the second which has a high experience loading ($\cdot575$) has the highest loading on *membership of professional associations* ($\cdot800$). Among all secondary modern teachers (ASM) as a group (i.e., college trained and graduate) experience items loaded on ASM F 5 $\cdot750$ on *first post* and $\cdot802$ on *first school* and in ASM F 6 $\cdot769$ for *number of previous posts*.

Grammar and graduate teachers seem to offer two factors in which teaching experience loads highly. The graduates throw up two distinct factors, whereas the grammar group has one factor relating to first and present post whilst the other high loading, on first school, appears in Common Factor 2 on teaching practice and college attainment.

Such a wide diversity of loadings might be expected in the light of the variety of teaching experience available, especially as some branches of teaching do not offer the same opportunities. Evaluation of the post, present or initial, is based on the responsibility and salary it commands. In primary schools it is almost impossible for a teacher to enter a post of responsibility on his first appointment; a situation far from unknown among science graduates starting their career in secondary schools. Graduates tend to gravitate to grammar and senior posts in secondary modern schools, whilst college trained teachers move into primary and secondary modern schools. This is not so in every instance but is a general trend. As such movements are basically due to types of training, the comparison of the progress made by a college trained teacher in a secondary modern school with that of his graduate colleague could not be based on a hypothesis of equal opportunity. Comparison of the responsibility of the initial post of all teachers is perhaps unrealistic when the opportunity for such responsibility is very restricted in one branch of the profession, i.e., the primary school, whereas it is quite rosy in other branches, e.g., the science graduate in a secondary modern school and the maths graduate in the girls' grammar school. For this reason the search for a *common* factor on the basis of the responsibility of the teacher's first post was perhaps not sound. When only graduate, grammar and all secondary modern teachers are considered, i.e., excluding college secondary modern and primary teachers, a common factor orientated towards first post can be seen, as in Table 6.

TABLE 6
A POSSIBLE FACTOR OF FIRST POST.

Variables	Graduate	Grammar	All Sec. Mod.
First Post	$\cdot785$	$\cdot783$	$\cdot750$
First School	$-\cdot077$	$\cdot264$	$\cdot802$
No. of Previous Posts	$-\cdot234$	$-\cdot165$	$\cdot127$
Responsibility of Present Post	$\cdot221$	$\cdot515$	$\cdot384$
Teaching/non-Teaching	$\cdot050$	$-\cdot204$	$-\cdot234$
Organizing Ability	$\cdot764$	$\cdot622$	$\cdot327$
Relationship with Children/Parents	$\cdot006$	$-\cdot074$	$-\cdot252$
Extra-Mural Activities	$\cdot420$	$\cdot088$	$\cdot103$
Head's Rating	$\cdot040$	$\cdot386$	$-\cdot052$
Total % Factor Variance	5.90	6.82	6.36
Factor Order.....	8	5	5

VII.—THE TEACHER'S FIRST POST.

After five years in the profession young teachers have had time to settle down and thus *present post* might be a common factor allowing for the differences in opportunities mentioned earlier. *Present post* appeared as a single factor for the primary teachers only, and then it was the seventh of the eight factors extracted accounting for only 5.82 per cent. of the variance. Loadings on *first post* in other teaching groups seemed to be moderate to low and spread over seven factors. An outstanding feature of all these factors which seems to relate to the first appointment of a newly qualified teacher (and, for that matter, the present status) is the absence of any loading from items based on work done on the training course. Neither *theory of education* nor *teaching practice assessment* appear to weigh in the balance of whether a primary school teacher goes to infant, junior or all-age school. This might be due to the very selective training received at college which prepared teachers of a wide range of ability for each category of school. However, a similar situation occurs in secondary teaching when the colleges do not prepare specifically for secondary modern, comprehensive, technical or grammar. This situation is re-affirmed among graduates and grammar school teachers. Whatever the reasons, *training course performance* appears of little significance to the type of school to which the teacher is first appointed (within his training dichotomy of secondary or primary).

What does determine the type of school to which a teacher goes to on his first appointment? Among graduates the demand for scientists presumably explains the grammar school appointment of more teachers from the science faculties than from the arts or social sciences (loading of $-.372$ on *faculty*). *Degree class* did not seem to be very important. Among the teachers in the secondary modern schools *qualifications on entry* to college appeared to bear some relation to *first appointment* but with nothing like the importance that this item had in the primary school. Once again it was noticeable that *qualifications on application for admission* to a training course had a higher loading than the qualifications the student possessed on *actual entry* to college, i.e., $.656$ compared with $.605$.

Among the teaching experience items the loading of *responsibility of present post* decreased from the grammar school through to the primary school. It is tempting to suggest that the better the first school the better the present appointment is likely to be and to link this with graduate status on the basis of the loadings of all secondary modern teachers ($.384$) and of all college trained secondary modern teachers ($.222$). However, the insignificant loading on the graduate group as such would suggest that such a tendency is one between groups rather than within groups. Among the graduates the degree of *responsibility of the present post* bore no relation to *first appointment* but among the secondary school groups, which included graduates, the degree of responsibility of the first appointment seemed important. It was least in the grammar schools, and this might be expected as the opportunities for responsible or senior positions in grammar schools for teachers on their first appointment after training must be relatively few, and, in the light of the *faculty* loading, occur in the science departments. Of the teachers now in grammar schools responsibility of *first post* seemed more important than type of school, or for that matter, than responsibility of *present post*. Thus, it would seem that a significant number of teachers from this sample who had responsible posts in grammar schools in 1960 started their teaching careers with responsible posts in secondary modern, comprehensive or technical schools. In the secondary modern school the

tendency for the *first post* appointment to relate to *present position* existed but to a lesser degree than the grammar, but was still rather more than in the primary school. The better the first school the fewer resultant changes of appointment were made by the teacher. This was particularly noticeable among the graduates and the college trained secondary modern teachers. Presumably graduates are happiest in grammar schools and non-graduates who are appointed to grammar or comprehensive schools do not wish to move. The trend exists at a much lower level in grammar schools as a whole ($-.165$), is insignificant at primary schools ($-.031$), but interestingly has a low positive loading in secondary modern schools as a whole ($.127$).

Graduate teachers with the better first appointments appeared to be less concerned with course work leading to *further qualifications* and particularly such courses as were considered of a professional flavour ($-.405$). The primary school teachers with the better first schools seemed to have followed courses of personal interest to themselves. Apart from this the decision to attend courses appeared independent of type of first appointment.

Membership of *professional organizations* seemed tied in with the *first appointment* only among the primary teachers and possibly the graduates. *Academic organizations* linked positively with *first appointment* in the college trained secondary modern teachers and the primary teachers ($.482$ and $.371$ respectively). There was a suggestion that the better the first appointment of the graduate the less inclined he was to join an academic association.

The better the *first appointment* the fewer the *causes of dissatisfaction* though this was not a marked trend, and the reversal of this trend in grammar schools is interesting, perhaps reflecting the dissatisfaction felt by the non-graduate in the grammar school who frequently commented on the status difference between graduate and non-graduate in the profession.

There appeared to be a steady descending loading on the headmaster's estimate of the teacher's *organizing ability* as one moved from graduate to primary (graduate $.764$, grammar $.622$, secondary modern $.327$, college trained secondary modern $.396$ and primary $-.260$). It would appear that those who had the better first appointments were considered to be better organizers if they were in grammar schools, but inferior if they were now in primary schools (e.g., previously been in secondary schools).

Of the head teacher's ratings only those in the grammar school loaded significantly on this factor and it seems that the better the *present post* of the individual and the more responsible his *first post* the more highly thought of he was.

VIII.—MEMBERSHIP OF PROFESSIONAL ORGANIZATIONS.

Common to all secondary groups there appeared a factor with loadings on *membership of professional organizations* (Table 7). Apart from the absence of any similar factor in the primary school this membership of *professional* rather than *academic* organizations showed only two items loading consistently throughout the secondary groups, namely those of the degree of responsibility of the teachers' *present post* which had a low positive relation, being much higher ($.417$) among graduate teachers than in the other groups, and *qualifications upon entry* to the secondary school training course.

Interesting differences between the groups in various loadings on this factor occur, as the moderate loading ($.575$) of *number of previous posts* among the college trained secondary modern teachers and the high loading ($.681$) of

number of causes of dissatisfaction among the graduates. It would seem that the graduates differ from the other groups on the issue of satisfaction and dissatisfaction as these appear as bi-polar variables on factor 4 for grammar, secondary modern and primary teachers and it is only in the graduate group that these appear with different factor structure. *Number of causes of dissatisfaction* is linked moderately ($-.449$) with satisfaction in factor 4 but has a high loading ($.681$) in factor 8 when associated with membership of *professional*, as distinct from *academic*, organizations. It also seems that more science graduates and graduates with some responsibility have more *causes of dissatisfaction* and are members of professional organizations, with a significant proportion being members of *academic bodies* as well.

TABLE 7
FACTOR 7—MEMBERSHIP OF PROFESSIONAL ORGANIZATIONS.

Variables	Graduate	Grammar	All Sec. Mod.	Non-grad. Sec. Mod.
Professional Organizations613	.693	.675	.800
Academic Organizations432	.010	.151	.036
Entry Qualifications	($-.125$)*	.510	.777	.250
Faculty	$-.465$			
Principles of Education238	$-.169$.061
Teaching Practice216	.178	$-.037$	$-.156$
Average Practice134	.313	.032	.236
Average Essay324			
Responsibility of Present Post417	.263	.279	.220
No. of Previous Posts	$-.050$	$-.261$.060	.575
Teaching/non-Teaching038	.108	.189	.635
Professional Courses293	$-.331$.093	.312
Personal Courses	$-.058$	$-.109$	$-.362$.058
Degree of Satisfaction	$-.200$	$-.050$	$-.002$.079
No. Causes of Dissatisfaction681	$-.274$.104	.150
Personality086	.071	.019	$-.247$
Organizing Ability251	.017	.198	.222
Relationship with Head and Staff ..	.119	.157	$-.204$	$-.236$
Head's Rating105	$-.260$.173	.193
Total % Factor Variance ..	6.35	5.21	5.64	7.44
Factor Order	7	8	7	5

* Degree Class.

IX.—ATTENDANCE AT COURSES.

Differences in loading on the *attendance at courses* are interesting. It seems that attendance at *courses leading to higher qualifications* does not correlate with *membership of professional organizations*—whatever the incentive to increase the status of the secondary school teacher given by the professional associations it is not in the direction of improving qualifications. It also seems that teachers who follow further *professional courses* tend to join *professional organizations* if they are graduates or if they are secondary modern teachers trained at college but that the reverse is the case for the graduate group. Among the secondary modern teachers as a whole *membership of professional associations* does not seem linked to professional courses at all, rather being negatively tied to courses taken for personal interest. Summarising, in the grammar and secondary modern teaching groups *attendance at courses* tends to reduce *membership of professional bodies*, whereas for the graduate and the college trained secondary modern teacher such attendance seems to stimulate membership.

TABLE 8

SUMMARY OF VARIABLE LOADINGS ON FACTOR 8.

Variables	Graduate	Gram-mar	All Sec. Mod.	Non-grad. Sec. Mod.	Primary
Age	-.624	-.557	-.526	-.609	-.245
Personal Courses	-.119	.602	.515	-.160	.668
Entry Qualifications175	.055	.623	-.057
Faculty221				
Principles of Education330	-.034	-.252	.084
Education Practice	-.345	.091	.170	.031	-.108
Average Practice	-.005	.454	-.123	.025	-.010
Responsibility of Present Post.	-.389	-.240	.036	.248	-.065
No. of Previous Posts	-.260	-.356	-.022	.108	.025
Professional Organizations024	-.084	-.025	.152	-.312
Professional Courses	-.017	-.063	-.026	.036	-.256
Academic Organizations	-.098	-.106	.553	.028	.102
Degree of Satisfaction025	.028	.041	-.022	-.214
Four Judges' Rating128	.017	.163	.018	-.275
Cognitive Ability	-.390	-.154	-.419	.024	-.238
Organizing Ability035	-.077	-.258	-.012	-.006
Relationship with Head and Staff ..	.314	-.037	-.094	-.134	.297
Extra-mural Activities	-.289	-.181	.020	.011	.285
Head's Rating201	.172	-.248	.342	-.311
Relationship with Children/Parents ..	.603	.227	-.102	-.107	.108
Total % Factor Variance	6.42	6.32	4.64	4.12	4.64
Factor Order	6	6	8	8	8

Individual group factors loading on age and to a lesser extent on attendance at personal courses showed fairly consistently and offered a small Common Factor with variable loadings summarised in Table 8. Attendance at personal courses after training appears to differ between the various teacher groups as to in which factor the high loading appears. In the graduates and the college secondary modern teachers this variable loads on Common Factor 3 (post-training courses) whereas for the grammar, all secondary modern and primary school teachers it loads on this eighth factor with low percentage variance. Whether or not the individual is *still teaching* seems to be a consistent trend can in the factor tending to favour those not now teaching and a similar trend can be seen for estimate of the teacher's *cognitive ability* where the direction is towards the low scores. For the graduates a very high loading (-.603) occurs with the item *relations with parents and children*, but it has a low loading (-.227) for grammar. *Head teacher's rating* appears to load significantly but not in the same direction for all groups of teachers. It would seem that this factor is associated with the *youth* of the teacher, his lack of *cognitive ability* and the tendency to have *left teaching*. Outstanding loadings are the good relations with children and their low *teaching practice mark* during training for the graduate and the *entry qualifications* of the college trained secondary modern teacher.

X.—SUMMARY AND CONCLUSIONS.

Questionnaire replies from 248 teachers were compared with performance at the final examination at the training college and the headmaster's report on the teacher five years after qualifying. The raw data were classified into teachers at different types of school, i.e., primary, secondary modern or grammar. A

further classification of the data was made into graduate and college trained teachers at the secondary modern school. Thus, five groups were available not all possessing the identical information (e.g., degree class). The numbers of variables available for the different groups varied from 42 to 35. The data were inter-correlated and subjected to principle component analysis. The first eight factors thus obtained were then rotated on the varimax principle. Four common factors emerged, *headmaster's opinion* of the teacher, *attainment on training course*, *attendance at post-qualifying courses* and *satisfaction with teaching* as a profession. The remaining factors were less consistent among all the groups. High loadings of items relating to the experience of the teachers appeared in at least two factors. *Membership of professional associations* seemed to be a factor common to secondary groups only and the high loading of age was the holding thread in the last somewhat heterogeneous factor. Individual groups had factors which did not fit any pattern among the other groups. For example, grammar fourth factor loaded on *principles of education* at college, *academic organizations* and *extramural activity*, and the college trained secondary modern sixth factor loaded on *cognitive* and *teaching ability* as assessed from the head teacher's report.

From the overall view a few things emerge. First, the minor feature of the higher loading of qualifications on *application* than qualifications on *admission* to training course on most of the factors (the exception being primary factor 7 classed in common factor 5). It seems that what a student acquires between application and entry actually decreases the relationship with seven of the eight factors isolated in this study.

Promotion as it is assessed by the degree of *responsibility of the teacher's present post* had its variance spread over seven of the eight common factors the exception being C.F.1, the *head teacher's estimate* of the teacher. This might suggest that the head's evaluation of a teacher bears little relation to that teacher's promotion—truly a peculiar situation. In the secondary school *present post* has loadings of between .281 and .441 on the factor of *attendance at post-qualifying courses*. *Present post* of all teachers has high loadings from .217 to .417 on factor 7, *membership of professional associations*. For graduates and grammar school teachers in general loadings occur on *training course performance* (factor 2) and factor 5 covering *first post* and *organizing ability*. The diversity of the ways to promotion and the lack of correlation between such promotion and the headmaster's opinion of the teacher suggests that the laws governing the teacher's progress in the profession are indeed complicated and mysterious.

A third interesting issue is the lack of agreement between *college assessment* and the *headmaster's reference* on the same individual five years later. On Common Factor 1, *headmaster's assessment* of the teacher, not one of the college assessments loaded greater than .18 and thirteen of the seventeen loadings are below .1. When Common Factor 2, *college attainment*, is examined the subsequent headmaster's rating of the teacher loaded —.015, .448, .246, .221, and —.120 from the graduate to the primary teacher. On the other hand, independent judges rating the teacher from the *headmaster's reference* only had one loading above .2. Thus agreement between headmaster and college seems quite small, achieving any worthwhile level for the grammar school teacher only.

Much discussion has centred on the importance of teacher training with considerable criticism of untrained teachers, be they graduates or not. It may be assumed that headmasters prefer trained teachers. However, it is evident

that the degree of success on the training course makes little difference to subsequent teaching performance as seen by the individual's head teacher, or even on the matter of the teacher's first appointment.

One important element in the teacher training course is that of *practice teaching*. Perhaps it might be expected that performance in the teaching situation when assessed by college lecturers might give an insight into the future performance of the teacher. Significant loadings on *practice* do not occur in Common Factor I dealing with *head teacher's reference*. Apart from a loading of .220 in the college secondary modern group for *attendance at courses* the only loading greater than .2 occurs with the graduate group: .422 on F.5—*satisfaction*, .216 on F.7—*membership of professional organizations* and —.389 on the heterogeneous factor 8 which loaded mainly on *youth and personal courses*. Factor 2, covering *training course attainment* and thus having high teaching practice loading, has *head teacher's rating* loading from .448 for graduates through to —.120 for primary school. Apart from the graduate and grammar school teacher it is difficult to see with what the professional training course is associated, certainly not with *promotion*, *head teacher's assessment* or *satisfaction in the profession*.

It may be that the colleges and the headmasters are using different criteria in assessing teaching ability and teachers in general. One might expect some criteria differences between college staff and headmasters. Similarly, the professional experience of each teacher must have varied considerably during the five years since he qualified and thus modified his behaviour. Hence, differences between the colleges' and headmasters' assessment were to be expected, but it is the magnitude of such differences that is surprising.

(Manuscript received 27th January, 1965.)

APPENDIX I

CONFIDENTIAL.

UNIVERSITY OF MANCHESTER SCHOOL OF EDUCATION RESEARCH COMMITTEE

Name

Address

Present Post

..... Salary.....

PREVIOUS PAID TEACHING POSTS :

Name of School (indicate the type of school if its title does not make it clear)	Status (e.g., Assistant ; Deputy Head)	Dates (give month and year)	
		From	To
FULL-TIME 1. <i>Pre-1955 (i.e., before training) if any :</i> 2. <i>Since 1955 :</i>			
PART-TIME (if any) :			

SPECIAL RESPONSIBILITY :

If any of the above posts involved special responsibilities (either with or without extra salary), please give details below :

School	Post	Date	Salary addition (if any)

COURSES ATTENDED SINCE 1955 :

If you have attended (voluntarily) any educational courses, no matter how brief, please give details below :

Date	Course	Organised by	Duration (No. of Meetings or Lectures)

MEMBERSHIP OF EDUCATIONAL SOCIETIES, UNIONS OR ASSOCIATIONS :

If, as well as being a member, you have at any time been a committee member, or held any office, please give details.

<i>Name of Society or Association</i>	<i>Office held (if any)</i>	<i>From</i>	<i>To</i>

ESTIMATE OF PRESENT LEVEL OF PROFESSIONAL SATISFACTION :

How *satisfying* do you find your job as a teacher ? Try to be as honest and frank as you can in your response. Put a cross in *one* of the squares.

☐

*Fully
satisfying.*

☐

*Satisfying on
the whole, but
not fully so.*

☐

*Moderately
satisfying.*

☐

*Some satisfac-
tion, but not a
great deal.*

☐

*Most
unsatisfying.*

My chief sources of dissatisfaction (if any) are :

ADDITIONAL NOTES OR COMMENTS (if any) :

Please add any additional information which you think is relevant to your development as a teacher since leaving college.

Will you give us permission to ask your present Head Teacher, *in confidence*, for a reference ? Under no circumstances will this be seen by anyone not directly concerned in this research. If you agree, we will write direct to your Head.

Name of Head Teacher

School Address

.....

VARIMAX ROTATION RESULTS—PART B

No.	ALL SECONDARY MODERN TEACHERS								NON-GRADUATE SEC. MOD. TEACHERS								PRIMARY SCHOOL TEACHERS																	
	1	2	3	4	5	6	7	8	HH	1	2	3	4	5	6	7	8	HH	1	2	3	4	5	6	7	8	HH							
1	026	555	-152	345	075	048	-013	-526	74	-142	276	-015	455	-042	239	-110	-609	75	-116	254	-003	-342	369	-133	377	-245	55							
2	034	-137	047	-101	237	187	777	055	73	067	-113	125	-310	250	-078	168	623	61	-037	299	183	655	-401	-213	-230	-144	74							
3																			032	203	160	605	-371	-208	-367	-057	83							
4																																		
5	-050	177	814	-221	053	-040	-169	-034	78	020	236	796	-192	061	229	059	-252	85	024	929	023	-017	-037	076	-012	084	88							
6	099	154	725	118	150	123	-037	170	62	179	220	783	120	-156	-075	088	031	75	109	736	187	-131	153	-030	102	-108	65							
7	036	170	817	-123	-118	-084	032	-123	75	-109	102	796	-129	238	276	119	025	82	038	902	082	124	-035	-029	024	-010	84							
8																																		
9																																		
10																																		
11	219	281	024	296	384	-471	279	036	66	039	441	-091	490	220	263	222	248	67	102	059	-067	217	-014	-042	776	-065	67							
12	100	032	-139	028	750	-181	-144	014	65	-048	-128	126	145	-113	068	768	049	66	-071	046	-015	676	-050	-153	092	111	51							
13	-055	-183	131	035	802	027	101	-047	71	-150	189	-071	145	575	097	-472	108	66	-129	-250	015	-031	-793	-033	213	025	76							
14	-131	-025	-064	016	127	-769	060	-022	63	056	222	028	147	635	-223	-056	-327	64	020	-159	183	-069	476	-138	198	-144	37							
15	361	370	-125	135	-234	-209	189	-078	44	-124	753	084	-008	-081	219	-082	020	65	-043	158	866	-020	042	-096	-088	107	81							
16	-038	817	157	-058	-077	013	009	-020	70	051	856	140	086	053	-014	-078	-011	77	-012	036	858	074	093	-043	009	172	78							
17	144	841	176	028	034	-046	028	171	79	-116	574	125	066	312	242	-115	036	53	162	226	535	076	058	-066	483	-256	68							
18	121	510	216	-228	-074	-441	093	-026	58	331	516	-080	-236	058	-148	-038	-160	49	037	-023	234	207	-010	-203	-146	668	61							
19	140	358	-112	-183	-157	-175	-362	515	64	126	-049	041	-104	800	142	021	152	71	-127	068	354	326	139	203	063	-312	41							
20	230	150	-025	-107	-079	-197	675	-025	59	041	574	137	-210	036	-259	371	028	60	-054	-182	148	482	263	-040	259	102	44							
21	-037	492	121	063	-011	292	151	553	68	155	034	-019	811	079	-005	077	-022	70	048	081	-025	-097	193	803	-033	-214	75							
22	070	132	074	798	041	-104	-002	041	68	-161	420	-082	-655	150	152	019	089	69	-071	041	134	113	246	-773	045	-016	70							
23	-040	292	117	-709	-017	-110	104	031	63	782	228	182	281	-003	-147	043	018	80	853	-086	-018	-023	173	021	-118	-275	86							
24	805	154	163	276	-018	009	111	163	81	753	-427	-098	-068	-247	004	-015	-012	82	850	051	-064	-089	-117	-036	019	105	76							
25	769	-274	072	012	017	324	019	-037	78	216	103	116	-083	-001	830	-006	024	77	348	144	-186	-111	-321	-325	-022	-238	45							
26	353	185	333	-164	100	027	011	-419	48	484	057	263	-260	010	-021	-142	-006	39	484	057	263	-260	010	-021	-142	-006	39							
27	628	148	067	-090	327	-109	198	-258	65	845	065	-003	037	014	102	090	-122	75	845	065	-003	037	014	102	090	-122	75							
28	771	078	279	150	016	-229	-073	140	78	729	142	028	110	-063	126	079	297	68	729	142	028	110	-063	126	079	297	68							
29	678	-149	-322	-068	132	196	-204	-094	70	698	-006	-103	-200	124	-050	289	108	65	698	-006	-103	-200	124	-050	289	108	65							
30	595	-005	032	039	-252	033	183	-102	47	360	041	117	266	469	156	215	285	59	360	041	117	266	469	156	215	285	59							
31	675	184	-074	-130	103	-045	035	020	53	959	100	008	-031	061	023	146	134	97	959	100	008	-031	061	023	146	134	97							
32	980	032	-020	-033	019	055	015	-104	98	646	-120	-016	126	-265	-111	-243	-311	69	646	-120	-016	126	-265	-111	-243	-311	69							
33																																		
34	693	170	246	300	-052	-231	173	-248	81																									
%	19	2910	-86	8	86	6	-46	6	-36	5	-68	5	-64	4	-64				18	29	9	68	8	-07	7	50	7	-07	5	-96	5	-82	4	-64

A NEW SCALE FOR PERSONALITY MEASUREMENTS IN CHILDREN

By SYBIL B. G. EYSENCK.

(*Institute of Psychiatry, University of London*)

SUMMARY. A brief description is given in this paper of the Junior E.P.I.: a questionnaire for use with children from the ages of 7 to 16. The questionnaire was developed by extensive factor analyses carried out separately for the various age groups considered, and detailed norms are available for boys and girls separately at each age. The questionnaire contains scales for the measurement of neuroticism or emotionality, extraversion/introversion, and a lie scale. The intercorrelations of these scales are given and there is a discussion of the relations observed between the temperamental traits measured on the one hand, and sex and age on the other.

I.—INTRODUCTION.

TEACHERS and educational psychologists have, in the past, concentrated their efforts almost exclusively on the cognitive field, at least in as far as normal children are concerned; even when children are sent to a child guidance clinic, intelligence tests figure more prominently in the investigation conducted by the educational or clinical psychologist than do measures of personality. One of the reasons for this state of affairs has possibly been the absence of properly developed questionnaires in this field, and the known lack of reliability and validity of projective techniques. Two questionnaires for children are in use at present, but there are objections to both. One is the Cattell IPAT scale (Porter and Cattell, 1960), the other the Junior M.P.I. (Furneaux and Gibson, 1961). In both cases it should be noted that there have never been conducted proper item analyses and factor analyses for all the questions used, and for the various age groups separately; it is suggested that such analyses are a requisite minimum qualification for a properly constructed scale. Furthermore, both scales are afflicted with American-type wording (the Junior M.P.I. was adapted from the Pintner Scale which, like the Cattell scale, was made and standardized in America). For these reasons it was considered desirable to create a new personality scale for children, which was named the Junior E.P.I. to denote its derivation from and resemblance to the Eysenck Personality Inventory, which is an improved version of the M.P.I. (Eysenck and Eysenck, 1964).

The scale was designed to measure the two major personality variables of neuroticism (stability/emotionality) and extraversion/introversion. The inventory also contains a lie scale for the detection of faking. The construction of the 108-item scale was begun by carefully selecting, adapting or re-writing items contained in the adult version of the E.P.I., and adding further items. These scales were administered to school children, mostly from the Rotherham area but also from various London schools. Their ages ranged from 7 to 16, inclusive, and both boys and girls were tested. It should be noted that the older children, i.e., those in the 15 to 16-year age groups are not as good a sample of the general population as are the younger children because some children in these age groups elect to leave school, so that those who remain are self-selected. We made up the numbers by including 16-year-olds from youth clubs and some apprentices in industry who had left school. Since the 16-year-old group is, by necessity, so heterogeneous it may be as well to interpret results for this age with some caution.

These 108 items were then intercorrelated separately for each age group so that in all nine correlation matrices were computed. The total number of children involved in this part of the work was 6,760. Factor analyses were carried out on all the matrices, except those of the 16-year-old children where the number was too small to make this worthwhile. The method of principal components was used and in each case the first two factors extracted contributed much the greatest part of the variance, and in addition could readily be identified as neuroticism and extraversion, respectively, by reference to those items known from previous work to measure these two dimensions. Table 1 gives the latent roots for the various groups of the N and E factors. It will be seen that there is no systematic change with age as far as the N factor is concerned but that there is a fairly steady increase in the size of the latent roots

TABLE 1
LATENT ROOTS.

n	Age	N	E	L
491	7	9.615	3.573	2.718
580	8	9.320	3.560	3.023
761	9	8.447	3.834	3.340
844	10	8.684	3.919	3.300
989	11	9.596	4.188	3.140
747	12	9.617	4.295	3.681
804	13	9.209	5.453	2.974
786	14	8.309	4.927	3.077
527	15	8.914	6.615	4.188

with age as far as E is concerned, with the 7 to 8-year-olds having relatively low latent roots, the 13, 14 and 15-year-olds having relatively high latent roots, and with the other groups intermediate. These results suggest that neuroticism can be measured adequately at all these age levels but that as far as extraversion is concerned either this is a personality dimension which is not clearly emerging until the age of 9 or 10, or its measurement by means of questionnaires presents unusual difficulty with the youngest age groups.

A lie scale of sixteen items was constructed by again adapting and re-writing items from the senior E.P.I. and adding others; this scale was given to a further set of school children not previously tested, amounting in all to 2,777. Intercorrelations were computed between these items for all age groups separately and nine factor analyses carried out, again omitting the 16-year-old groups. One factor only was extracted, loading on all the items and contributing latent roots which are indicated in Table 1. There appears to be a slight increase in the size of the latent roots with age but this is probably too small to be of any great significance.

On the basis of these factor analyses, sixty suitable items were chosen for the final inventory of which twenty-four measured E, twenty-four N and twelve constituted the L scale. The choice was based on the loadings of the items for their respective factors, and their lack of loading on other factors. The correlations between the E and N scales for the different age groups are given in the Manual (S.B.G. Eysenck, 1965); they are positive in three cases, negative in the rest, and on the whole the figures suggest a slight negative correlation between neuroticism and extraversion which, however, is too small to be of any practical importance.

Standardisation data are also given in the Manual. They show certain interesting changes in scores on both extraversion and neuroticism with age, and also some interesting sex differences. Figure 1 gives the scores for extraversion. The trends shown in this Figure, i.e., the regular increase for both sexes up to the age of 14 or so, and a higher mean score at all levels for boys, were shown by analysis of variance to be significant beyond the .001 level.

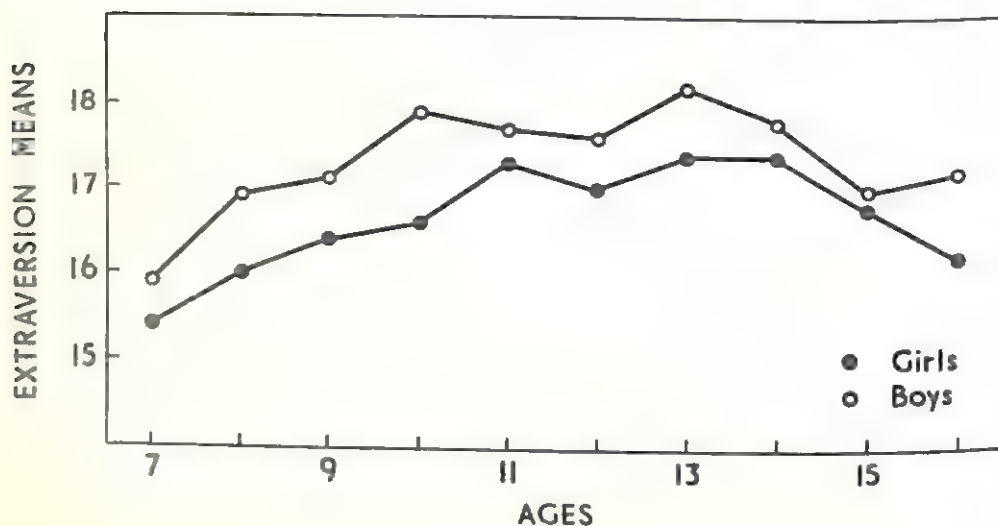


FIGURE 1

The sex difference observed here confirms similar findings with adult populations where also males tend to obtain higher scores than females.

Figure 2 gives a similar picture for neuroticism scores. It will be seen that while both groups start out on roughly the same level, girls show a consistent increase, boys, if anything, a slight decline. Analysis of variance demonstrated that sex, age and the sex by age interaction were all significant beyond the .001 level. It thus appears to be a fact that girls become more emotionally unstable with increasing age, whereas boys remain at much the same level, or even become somewhat more stable. (It is assumed, of course, that the questionnaire scores are indicative of true life behaviour).

As regards the lie scale there is a marked shift from high scores in the younger children to low scores in the older children. This shift is quite monotonic and is best shown in Figure 3 which gives lie scores in percentages for the oldest and the youngest groups tested, respectively. This finding is in good agreement with the known lying behaviour of children (McFarlane, *et al.*, 1954).

The split-half reliabilities of the scales indicate that they are reasonable for group comparisons at the younger ages, and possibly for individual testing at the older ages. For neuroticism the values range between .8 and .9, and there is no detectable change with age. For E there is an increase with age from values around .65 for the youngest groups to values approaching .9 for the oldest groups. The lie scale, as might have been expected from the fact that it is rather short, shows the lowest reliabilities ranging from .41 for the youngest groups to .78 for the older ones.

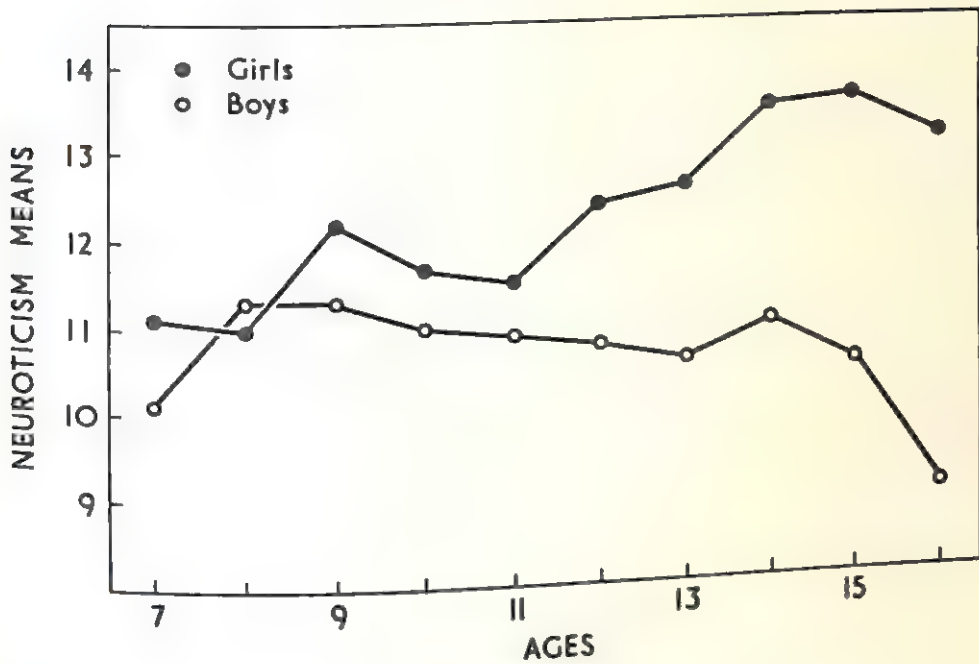


FIGURE 2

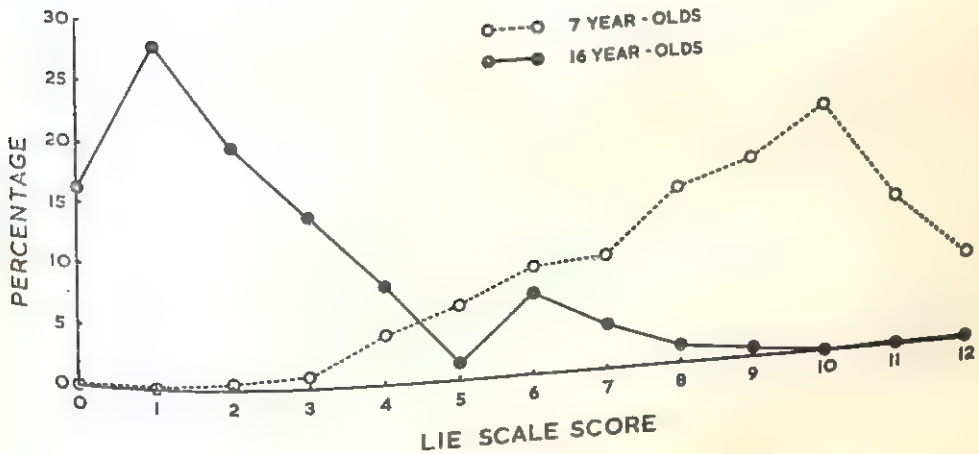


FIGURE 3

Test/retest reliabilities are available for additional samples of 1,056 boys and 1,074 girls. Typically these range between .7 and .8, with isolated values up to above .9. There is a tendency for test/retest reliabilities to increase with age for the E scale, and to a slight extent for the N scale as well; there seems to be no such tendency for the L scale. Altogether test/retest reliabilities are somewhat lower than split half reliabilities.

It has already been mentioned that correlations between E and N are rather slight, averaging above $-.15$; this does not appear to vary very much with age or sex, and is, of course, too small to be of any practical importance. There is a slight negative correlation for the older children between E and L of approximately the same size, which is in good conformity with data found in adults. Correlations between N and L are rather larger, averaging above $-.3$ and in some cases exceeding $-.4$. The size of the correlation appears to be largely independent of age, and the data suggest that children with high lie scores will tend to have low neuroticism scores, possibly due to conscious attempts at faking. There is too little direct evidence as yet to establish the usefulness of attempts to control for this by any known statistical procedure, but further experimental work, particularly using criterion groups of known neurotic children, may indicate optimal ways of controlling attempts to fake.

The personality theory on which this test is based suggests that E and N are independent of intelligence, and work with adults has usually substantiated this hypothesis. It seemed desirable to obtain additional information for children, and data were obtained on 373 girls and 334 boys in all. The children had been given the Junior E.P.I. as well as two verbal intelligence tests; ages represented were from 11 to 16, but the number of cases in the 16-year-old group was too small to give reliable data. The data show no evidence of any correlation between E and I.Q.; they show a slight tendency for N to be negatively correlated with I.Q., but the numerical values of the coefficients are small and not always in the same direction. There is a definite tendency for lie scale scores to be negatively correlated with intelligence although here again an overall value of the correlation is only between $.2$ and $.3$. It had already been found in some unpublished work with an E.P.I. scale specially prepared for subnormal subjects that these give much higher lie scores than normal subjects so that this association must be regarded as being very likely a true picture of the situation.

Too little is known as yet about the validity of the Junior E.P.I. to make any claims for its use, other than as an instrument for experimentation. Two-hundred-and-twenty-nine child guidance clinic subjects have been tested and rated with respect to the extraverted or introverted nature of their symptoms. It was found that the group as a whole was very significantly above the standardisation group with respect to neuroticism, and that there was a very significant difference with respect to E between children showing extraverted symptoms and those showing introverted symptoms. It seems probable, therefore, that the scale may have some validity in connection with clinical investigations and it seems not unreasonable to expect that a similar validity will attend measurements in normal children.

We may perhaps end by pointing out some of the areas of research where this scale may be most useful.

(1) It has been demonstrated that intelligence test scores may show a curvilinear relationship to neuroticism and may, therefore, require correction to give optimal forecasting efficiency as far as school achievement and other criteria are concerned (Lynn and Gordon, 1961). This is an important area of research in view of the importance which intelligence tests and other anxiety producing examinations have in the life of children. Indeed, the whole structure of intellect may be different in persons of different personality (Eysenck and White, 1964).

(2) School and university achievement has been found to be related to introversion and neuroticism, as well as to intelligence, and much further work is needed to put these findings on a secure basis and disentangle the causal effects which may be responsible (Furneau, 1962).

(3) Experiments in educational psychology have usually consigned personality differences to the error variance rather than investigate them directly. This seems a strategic error which may have far-reaching consequences. To quote but one example, Thompson and Hunnicutt (1944) have shown that the work of introverts improves with praise, that of extraverts with blame. If they had simply treated these personality differences as part of the error variance they would have come to quite erroneous conclusions and would have missed a finding which may be of crucial significance for teachers and educationalists alike.

(4) Among the many investigations which may benefit from closer attention to personality variables are studies using teaching machines. It is unlikely that the conclusions usually drawn in published papers applied equally to all children: it may be suggested with diffidence that extraverted and introverted children, or stable and unstable children, may show quite different and even contradictory reactions to identical conditions of testing.

(5) The same may be said of experimental variants in classroom procedure. There is already evidence that extraverted children accumulate reactive inhibition much more quickly than do introverted children and would, therefore, benefit from repeated rest pauses, which are unnecessary for the introverts (Grassi, 1964). Theory suggests many other ways in which individual differences between children may cause them to respond differently to different classroom arrangements, but enough has been said already to indicate that few educational experiments can satisfactorily be designed without taking into account individual differences in personality. It is in this field rather than in the clinic that we believe the major importance and usefulness of the Junior E.P.I. to lie.

II.—REFERENCES.

- EYSENCK, H. J., and EYSENCK, S. B. G. (1964). *The Eysenck Personality Inventory*. London: Univ. of London Press.
- EYSENCK, H. J., and WHITE, P. O. (1964). Personality and the measurement of intelligence. *Brit. J. Educ. Psychol.*, 34, 197-202.
- EYSENCK, S. B. G. (1965). *The Junior Eysenck Personality Inventory*. London: Univ. of London Press.
- FURNEAU, W. D. (1962). The psychologist and the university. *Universities Quarterly*.
- FURNEAU, W. D., and GIBSON, H. B. (1961). A children's personality inventory designed to measure neuroticism and extraversion. *Brit. J. Educ. Psychol.*, 31, 204-207.
- GRASSI, J. (1964). *Experimental Investigation of Cortical Inhibition in Children*. London: Unpublished Ph.D. thesis.
- LYNN, R., and GORDON, I. E. (1961). The relation of neuroticism and extraversion to intelligence and educational attainment. *Brit. J. Educ. Psychol.*, 31, 194-203.
- MCFARLANE, J., ALLEN, L., and HONZIK, M. (1954). A developmental study of the behavior problems of normal children. *Univ. Calif. Publ. in Child Developm.* Vol. 2, Univ. of Calif. Press.
- PORTER, R. B., and CATTELL, R. B. (1960). *The IPAT Children's Personality Questionnaire*. Champagne, Ill.: Inst. Person Abil. Testing.
- THOMPSON, G. G., and HUNNICUTT, C. W. (1944). The effect of repeated praise or blame on the work achievement of 'introverts' and 'extraverts.' *J. Educ. Psychol.*, 35, 257-266.

(Manuscript received 6th April, 1965.)

RESEARCH NOTES

FACTORIAL STUDIES OF PERSONALITY AND THEIR BEARING ON THE WORK OF THE TEACHER

By CYRIL BURT

(*University College, London*)

I.—EARLY INVESTIGATIONS.

In the February issue of this *Journal*, Dr. Adcock (1965) has undertaken an instructive comparison of the various concepts put forward by Professor Cattell and Professor Eysenck in their studies of personality. If I understand him aright, the practical corollary that he wishes to draw is that, in spite of obvious differences in detail and the rather drastic criticisms already advanced by clinical psychologists (this *Journal*, 1946), the agreement between the two—one working at an American university and the other at a British mental hospital—is sufficiently impressive for educationists and psychologists to accept their main conclusions without demur. "The most striking fact," he writes, "is that there are probably no findings which cannot be reconciled."

A sceptical reader, however, might offer an alternative explanation for the basic similarity of their views. Although each has worked in a different field and has developed his analytic procedures along somewhat divergent lines, both commenced their factorial studies at University College, London, and so started with much the same hypotheses. And this, he might argue, is the explanation for the resemblances between their statistical methods, and for their apparent agreement over the two or three main 'personality factors.' Moreover, both of them have carried out their researches on adults: and adult subjects brought up in different countries—especially those encountered in colleges and mental hospitals, respectively—tend to develop distinctive peculiarities in those particular characteristics which are most open to observation. That of itself might account for the minor modifications which they have introduced into their schemes, and to which Dr. Adcock has drawn our attention.

Readers of this *Journal* will doubtless be chiefly interested in the variations in personality shown by children; and it was principally with children of school age that my own early work was concerned. The aim was threefold: first, to discover whether there is any objective evidence to verify the widespread notion that differences in temperament and motivation are partly determined by innate or hereditary influences; secondly, to ascertain whether the early observation of such differences might help to indicate which children were likely to develop neurotic or delinquent tendencies—questions that were very much to the fore during the first world war; and thirdly, to throw light on a problem which the practical teacher is continually raising—why do so many pupils of normal or even super-normal ability fail to achieve the educational standards that one would naturally expect from their intellectual capacities? The investigations which I and my co-workers (chiefly teachers) carried out, and the results obtained, are scattered in a number of different reports and papers, some no longer accessible. Hence, it may perhaps be helpful to begin with a brief summary of the conclusions that were drawn.

The term 'personality' is nowadays used in a somewhat narrow and specialized sense to denote "the distinctive way in which any given individual's non-cognitive or dynamic tendencies are organized," i.e., the various affective, conative, and emotional propensities which chiefly determine his interests, motives, preferences,

and indeed, his whole social and personal behaviour. The descriptive terms currently employed by teachers, educationists, and psychologists, as well as by biographers, historians, and novelists, are so numerous, and evince so much over-lapping, that it seemed desirable to follow the procedure already employed in cognitive investigations, and start by "reducing them to a small number of relatively independent tendencies, capable of serving as principles of classification." Wundt (1910) had already endeavoured to analyse the more composite feelings and emotions into *Faktoren oder Komponenten*, and concluded that: "the entire system can be regarded as a three-dimensional manifold, in which each dimension includes two opposite directions"; but his theoretical scheme seemed to most contemporary psychologists "so subjective and speculative" that Külpe in his textbook declared "we cannot accept any such classification of the feelings or emotions."

(a) *Normal Children*. In my first investigation (Burt, 1915), carried out partly with the assistance of Mr. R. C. Moore, the group selected for study consisted of 172 boys and girls, aged 9 to 11, selected as a fairly representative sample of the pupils attending the junior classes in the elementary school. Estimates were obtained for what McDougall held to be the "primary emotional tendencies," These particular traits were chosen partly because they were precisely defined and illustrated, and partly because, according to McDougall's biological assumptions, individual differences in these traits were themselves presumed to be innate. Eysenck (1952, p. 41) has criticized my selection as "thirty years out of date, and discredited in contemporary psychology." Cattell (1957, p. 536), on the other hand, maintains that this list of 'drives' is fully in accordance with current biological and psychological views, and has himself adopted it for some of his investigations. The estimates we obtained were based partly on first-hand observations under standardized conditions—carried out by teachers and ourselves, and partly on would-be objective tests—associative reactions, psychogalvanic reactions, apperception tests, inkblot tests, story tests, picture preferences, etc., all rather similar to what later writers have termed 'projective tests' (Burt and Moore, 1912, see also this *Journal*, XV, pp. 112-116).

The correlations between the final assessments for these 'primary emotional tendencies,' were then factorized by the method of 'simple summation,' and three statistically significant factors were obtained. There was first a general factor, analogous to the factor of general intelligence on the cognitive side, which I termed 'general emotionality' and tentatively identified with McDougall's 'common fund of emotional energy': the assessments for this factor showed a high but by no means perfect correlation with the teacher's assessments of each child's 'temperamental instability' (a popular phrase in the educational textbooks of those days). The two remaining factors were 'bipolar': the first divided the emotions into a 'sthenic' or 'ad-gressive' group and an 'asthenic' or 'ab-verse' group, i.e., those making for an active approach towards the environmental stimulus (assertiveness, curiosity, anger, sociability, sex, etc.) and those tending to inhibit such tendencies and making rather for withdrawal (fear, disgust, grief, submissiveness, etc.); the second divided them into a 'euphoric' and a 'dysphoric' group, i.e., emotions accompanied by a pleasurable or an unpleasurable feeling tone, respectively. The scheme was illustrated by a circular 'cosine clock-diagram,' in which the cosines of the arcs separating the several traits were proportional to the distances between adjacent pairs (the table of correlations is reproduced to illustrate factorial methods in this field in Holzinger and Harman, 1941, p. 176, and in Harman, 1960, p. 176). This classification of emotional propensities, I suggested, could also be used to classify persons into 'temperamental types'—type being regarded as a matter of degree. In particular, the dichotomous 'type-factors' appeared to provide a plausible basis for the traditional scheme of temperaments which has come down to us from Galen: the choleric is an active dysphoric; the melancholic an inactive dysphoric; the sanguine an active euphoric; the phlegmatic an inactive euphoric. The normal individual would, therefore, correspond with Galen's 'well-balanced' type—neither fantastically melancholy, too slowly phlegmatic, too lightly sanguine, or

too rashly choleric."* Suggestions as to the underlying physiological conditions (peculiarities of the autonomic and glandular systems) were also tentatively put forward.

To show that the factors thus obtained were not peculiar to the period of childhood, we carried out a parallel study with a group of young adults (chiefly students and members of a working men's club). Virtually the same three factors reappeared. In a series of later papers (Burt, 1917, 1937, 1938, and refs.) I gave brief case-histories illustrating the commoner types as they occurred among school children of both sexes, and endeavoured to show how the resulting differences in motivation affected the child's educational progress in school. At the same time, we tried to broaden the basis for the conclusions thus reached by including a wider variety of groups—normal, neurotic, and delinquent, and adopting a fuller list of observable traits, drawn in one investigation from the descriptive terms used by educational, psychological, and psychiatric writers, and in another from Roget's *Thesaurus*. As a result, a number of supplementary factors were also tentatively identified—in particular, narrow group factors corresponding with McDougall's 'primary emotions' or 'propensities,' and others underlying various acquired characteristics. The whole appeared to form a 'hierarchical system,' analogous to that obtained on factorizing cognitive traits, and for much the same reasons, namely, the inherent organization of the nervous system (cf. this *Journal*, 1949, Fig. 1, p. 103; Eysenck, 1953, p. 13; Cattell, 1957, p. 427). Correlations were also calculated between the estimates for temperamental types and physical types, and proved to be positive, though rather low (Burt, 1938, Table XI, p. 186).

Some confirmation for the original scheme was found in the fact that the distinction between 'sthenic' and 'asthenic' types corresponded in many ways with the distinction between 'extraverted' and 'introverted' types drawn by Jung (1923), between 'excitatory' and 'inhibitive' types drawn by Pavlov (1941), and between 'cyclothymic' and 'schizothymic' types drawn by Kretschmer (1926). Under slightly different names, similar distinctions had long ago been noted by James, who talked of 'explosive' and 'obstructed' types, and by Guthrie, who described 'restrained' and 'unrestrained' types among children. Out of this multitude of alternative labels we eventually decided on the Jungian terms 'extravert' and 'introvert' since they provided nouns as well as adjectives. Oddly enough, the distinction between the euphoric and the dysphoric temperaments seems rarely to have been mentioned by later psychologists, whether factorists or clinicians. The contrast between the two—the optimist and the pessimist, Autolycus's 'merry heart' and 'sad,' Milton's *L'Allegro* and *Il Penseroso*, Herbart's *Lustvolle Kinder* and *Unlustvolle Kinder*—seems commoner in general literature than in academic psychology. A survey of the whole evidence was summarized in a contribution to Reymert's *Symposium on Feelings and Emotions* (1950).

(b) *The Neurotic Child*. The relevance of these factorial theories to the study of neurotic and psychotic disorders in children I have discussed in some detail in the *Subnormal Mind* (1935) and more recently in my Maudsley Lecture (1954). A factorial analysis of neurotic cases revealed a clear distinction between two main

* BEN JOHNSON, *Cynthia's Revels*. Galen (court physician to the Emperor Marcus Aurelius and the undisputed authority on medicine for the next 1,400 years) was far more concerned with the medical implications of the theory of 'humours' than with the psychological. He recognized nine types in all, viz., in addition to the 'well-balanced temperament—a warm (active), a cold (inactive), a moist (comfortable), and a dry (uncomfortable) type, as well as the four more familiar 'mixed' types. The doctrine derives not from Hippocrates and the medical school of Cos, but from Empedocles and the rival Italian school. Those psychologists who ascribe it to Hippocrates have evidently not read the authentic Hippocratic tracts in the original Greek; the Hippocratic school rejected the Empedoclean doctrine as too speculative. To assume (as many psychologists apparently do) that because a treatise is to be found in the 'Hippocratic collection,' it, therefore, emanated from Hippocrates is rather like supposing that the books in the Bodleian library were all written by Bodley.

types—the 'asthenic' (e.g., neurasthenia, various anxiety-neuroses, phobias, and the like) and the 'sthenic' (the obsessional neuroses, compulsion neuroses, conversion-hysteria, anger neuroses, and certain forms of delinquency). As with delinquency, however, so with neurosis, "any actual nervous breakdown appears nearly always to be assignable, not to just one or two major factors, but to a wide variety, and usually to a multiplicity of converging and cumulative influences." We found it convenient to distinguish three main groups of causes—genetic, developmental, and precipitating. To determine the influence of genetic conditions, as distinct from post-natal, my co-workers and I adopted the same twofold approach as in the study of intellectual disabilities—comparing (a) children of different heredity who had been subjected to similar types of environmental stress, and (b) children of similar heredity but different post-natal histories. We found some evidence for the inheritance of emotional instability as a multifactorial trait; and it appeared that the really stable child rarely succumbed to nervous breakdown, however great the shock or strain. Generally speaking, however, the child's innate constitution tended to determine what type or form the nervous breakdown would take rather than the mere likelihood of any actual breakdown. During the recent war, the chief environmental disturbances were, first the shocks and strains of the air-raids (a factor we had also studied during the first world war), and secondly, the separation of the younger children from their parents see this *Journal*, X, pp. 9-15 and refs.). Different children reacted in very different ways: some, for example developed 'asthenic' disorders, particularly anxiety-states or phobias often accompanied by incontinence; others developed 'sthenic' disorders, such as neurotic compulsions and other obsessional states, and not infrequently some form of delinquency. In the case of the youngest evacuees—suffering, so to speak, from an artificially 'broken home'—there was often a delayed reaction: the more serious after-effects did not appear until puberty or adolescence. On examining the family histories we discovered, even among those who showed no signs of any emotional disorder, similar temperamental traits—'asthenic' or 'sthenic' as the case might be. The study of siblings and twins—particularly of monozygotic twins brought up separately—confirmed these conclusions. Thus, we found two pairs of a markedly 'asthenic type': in each case the child who lived in the country remained to serious anxiety neurosis, while the twin who had lived in the country remained perfectly normal, but showed the usual asthenic or introverted characteristics. Miss Conway found another pair, both of a markedly 'sthenic' type: in this case the Londoner succumbed to an obsessional neurosis, followed by a long spell of delinquency during puberty. Other investigators have reported similar observations (Lewis, 1935; Brown, 1942; Cohen, 1951; Rudin, 1953; and Shields, 1962). In peace time the strains and stresses of the school period—a broken home, an congenial school or teachers, overwork, the pressure of examinations—appear to operate on different individuals in much the same corresponding ways.

The factorial scheme that I have described has served as the basis for a 'schedule for case-records' published in several L.C.C. reports and later publications (e.g., Burt, 1937, p. 633), and has been used fairly widely both by teachers and at child guidance clinics. But it is without a doubt capable of considerable amplification and improvement. Let us, therefore, now inquire how far it needs to be modified in the light of the later researches by Eysenck and Cattell.

II.—EYSENCK'S FACTORIAL SCHEME.

To begin with, it may be noted that many of the differences between these two investigators spring from the fact that each is primarily interested in a different level of the general hierarchy—Eysenck in the broader group factors responsible for 'types,' Cattell in the more specialized factors forming 'source-traits.' In spite of their mutual criticisms, therefore, their views, as Adcock rightly argues, are complementary rather than antagonistic.

The work of Eysenck and his collaborators has been carried out mainly at the Maudsley Hospital; and, as Sir Aubrey Lewis notices in his preface, it was primarily "clinical categories that provided the starting-points for his investigations." His

first research (Eysenck, 1947), the results of which, he tells us, "underlie most of (his) subsequent work," was based on a study of 700 "male senile patients." Four factors were extracted: but two were rejected as devoid of significance, both statistical and psychological. "We are left," he says, "with two main factors or principles of classification—(1) a general Neuroticism and (2) a dichotomous division between Hysteria and Dysthymia." The former, he suggests, "comes close to the conception of the 'neurotic constitution';" the latter he regards as similar to the dichotomy between extraversion and introversion, aggressiveness and inhibiteness, and analogous distinctions reported "in a multitude of earlier studies." Apart from the change to a psychiatric terminology, these "two main factors" seem at first sight to correspond with those which my co-workers and I had found in nearly all our studies. In a footnote, however, Eysenck briefly indicates how his scheme differs from mine. "Pride of place," he writes, "might seem to go to Burt's pioneer investigation (1915) were it not for difficulties of interpretation: if we regard the introvert as the emotional person and the extravert as the unemotional, then Burt's general factor of emotionality might be identified with the extravert-introvert dichotomy." This I find rather puzzling. I had definitely identified my *second* factor with "the contrast between the aggressive or extraverted type and the inhibited or introverted," and I explicitly described the typical extraverts as "those whose emotions are unrestrained" (*loc. cit.*, p. 54); hence they could not possibly be regarded as "unemotional."

Dr. Adcock rightly asks why the first factor is called 'neuroticism' rather than (as he suggests) 'emotional reactivity' (which I myself would certainly prefer). The answer seems clear when we look at the list of traits for which the "patients" were assessed (Eysenck, 1947, Table 1). Few of the traits indicate simple emotional tendencies, and those with the highest saturation are specific neurotic symptoms—tremor, headache, dyspepsia, fainting fits, hypochondria, abnormal before illness, and (highest of all) 'badly organized personality.' In addition, Dr. Eysenck also includes conditions recorded in the case-histories which were liable to precipitate a neurotic breakdown—unsatisfactory home, abnormal parents, domestic problems, bombing, exposure, wartime separation, and the like. He argues, "with certain reservations," that the same "dimensions of personality" are to be found in normal members of the population. Here, however, I should be inclined to agree with one early reviewer that "the proposal to describe normal variations in terms drawn from psychopathology and to regard the healthy individual with strong emotions as suffering from 'neuroticism' is not only inapposite, but positively misleading."

In his second book (1952) Eysenck deals more fully with data obtained from normal people. He finds, he tells us, the same two factors as before. But, in his chapter on 'operational definitions' he now describes the first factor as 'a general factor of neuroticism or stability,' and renames the second 'extraversion.' He also proposes a third dimension of 'psychoticism.' Vernon (1964, p. 192) very pointedly asks: "What would be the characteristics of the extreme non-psychotic, and in what sense could the average or normal individual be thought of as half-psychotic?" As regards particular types of psychotic disorder (especially the commonest of all—schizophrenia) the evidence for the inheritance of a *specific* liability is so strong that few psychiatrists are willing to accept the notion of a *general* psychotic factor.

In his third book Eysenck (1953) presents his three-dimensional scheme as "a synthesis of the views of earlier writers," and vigorously defends the psychiatric labels. He again insists that his first factor is "clearly similar in conception to the general factor of 'intelligence,'" but criticizes my own interpretation still more sharply. He quotes an earlier research of mine on 483 normal and 328 neurotic children, based as before on emotional and temperamental assessments, in which I had maintained that, in spite of minor variations in the factor-saturations, the general factors were the same in both groups. "Since the most conspicuous difference between the neurotic and the normal is the emotional instability of the former, the selection of cases from the upper end of the emotional scale (so I argued) would naturally lower the size and alter the proportions of the saturations." Dr. Eysenck, however, observes that "one's faith in Burt's identification of his factors must be

considerably diminished by his failure to take account of the obvious discrepancies: it is difficult to see how he can identify the two factors when the correlation between their saturations is not significantly different from zero." But here Dr. Eysenck is mistaken. The amount of resemblance between two factors is measured by the ratio of the product-sum of their saturations, *taken as they stand*, to the geometrical mean of their squares, not by their correlation (i.e., the ratio based on their *deviations*); and measured in this way their resemblance is little short of 0.9.

In a recent symposium, Eysenck (1964) has again described his scheme, and illustrated it with a clock-diagram very similar to my own in which the relation between adjacent traits is shown by "the cosine of the angle of their separation." The two main factors are now described as (1) Stable *versus* Unstable and (2) Extraverted *versus* Introverted; the psychiatric designations are dropped. He points out how the results fit in with "Galen's system" (the four traditional temperaments), and notes the resemblances to "Wundt's dimensional view." "Cattell's most recent book," he observes, "shows a firm agreement with the system I (Eysenck, 1947) first put forward." The traits now included in his diagram would seem to bring this modified version of his scheme into fairly close accord with my own. One minor point perhaps deserves a passing comment. As I have elsewhere argued, the degree of a person's instability does not always correspond precisely with his degree of general emotionality. In a low grade defective even a moderate amount of emotionality is enough to render him unstable: on the other hand, as the biographies of some of our most eminent twentieth century statesmen suffice to show, exceptionally strong emotions can be integrated in the personality of a man of high intelligence, and so far from making him neurotic or unstable, serve merely to enhance his mental energy. Hence, a better estimate of a person's instability may be obtained by taking the ratio of his general emotionality to his general intelligence.

There is one further point (not mentioned by Adcock) on which Eysenck agrees both with Cattell and myself and on which we all three differ from the prevailing view—namely, the influence of the individual's genetic constitution—a matter, so I believe, deserving special attention in the study of children. But there are minor differences also. My own data, both for twins and for ordinary siblings, yield much lower correlations for emotional characteristics than for intelligence. Eysenck finds "as substantial an hereditary component in measured neuroticism as is commonly found with intelligence." On the other hand, Shields concludes from his investigations of twins at the Maudsley Hospital that what is inherited "is not so much stability or neurotic potential as their emotional responsiveness" and their degree of "extraversion—introversion": his correlations for both these qualities are likewise decidedly lower than for intelligence (Shields, 1962; Vernon, 1964, p. 188).

III.—CATTELL'S FACTORIAL SCHEME.

On turning to compare Cattell's factorial scheme with that of Eysenck, Adcock very properly starts with the question of choosing between (i) a general factor supplemented by bipolars or (ii) a miscellaneous set of correlated and positive group factors. "We would argue," he writes, "that it can be decided on grounds of parsimony." But there are still more important grounds than this. First of all, much depends on the nature of the traits: with cognitive abilities antithetical would prefer positive group factors to bipolar; with emotional tendencies antithetical descriptions seem more natural. Secondly, much must also depend on the investigator's purpose: if his primary object is classification, then a succession of dichotomies will often prove more effective; if, with Cattell, his interest lies in the causal influences, then a variety of co-ordinate group factors would seem almost inevitable. Adcock, however, notes that, whereas Cattell takes out a dozen or more factors, Eysenck at most takes out only four; and he supposes that this is because Eysenck "plays safe." But a matrix containing only 4 bipolar factors may yield a matrix of $2^4 = 16$ group factors; and with a complete hierarchical scheme a bipolar matrix of n factors may yield a system of $1 + 2 + 4 + \dots + 2^{n-1} = 2^n - 1$ overlapping 'broad' and 'narrow' group factors.

In the review already cited (Reymert, 1950, pp. 550f) there is a brief account of the similarities between Cattell's scheme of personality and my own; this, however, is based on his earlier book (1946). In his latest volume (1957) he himself indicates how his correlated 'first order factors' may be reduced to a few 'factors of the second order.' Of these, the first is designated 'Extraversion-*versus*-Introversion,' and the traits enumerated include most of those that characterize my first bipolar factor; the second is called 'Anxiety,' and for this I fully agree with Adcock, that we might well "read 'Emotional Reactivity'."

Here, however, I should like more especially to draw attention to a study which Adcock has apparently overlooked. For one of Cattell's most instructive researches covering some thirty-six traits, we are given his initial table of correlations as well as his factor loadings. Dr. Banks (1948) has carried out a complete refactorization of the table, with rotations, using the methods I myself had adopted. She obtains four significant factors—one general and three bipolar. The first she identifies with 'general emotionality'; the next—the first bipolar—is strongly weighted with traits accompanied by a pleasurable or unpleasurable feeling-tone; but since it also includes certain neurotic symptoms, she prefers the label 'euthymic *versus* dysthymic' to 'euphoric *versus* dysphoric': the second bipolar (she says) suggests a contrast between "a strong or sthenic attitude" and "a weak or asthenic"; but once again the psychiatric traits imply a broader contrast between the 'tough, hardened, firmly organized personality' and the 'feeble, undeveloped, ill-organized and inhibited type.' The last bipolar factor is definitely psychiatric; and suggests different contrasts for different groups. The merits which Dr. Banks claims for her procedure are that it is at once more objective and more speedily accomplished. Whereas Cattell's analysis "required 588 rotations and eighteen months' full-time work," hers required scarcely as many hours. Her rotation is based directly on the sign-pattern shown by the initial summational factors, and thus indicates the relation between the two alternative modes of classification. The final result is "a fairly systematic hierarchical scheme."

I think, therefore, that we may reasonably say that on all the major issues there is, underlying the apparent differences between all three of us, a large measure of agreement. The main points may be summed up as follows: (i) Though each personality is unique, it nevertheless exhibits in every individual much the same recognizable structure, namely, a hierarchical system, comprising a number of distinct motivational tendencies, partly innate and partly acquired; (ii) the most influential tendencies are the two so-called 'type-factors,' namely, (a) a general factor, which I termed 'emotionality,' and Dr. Adcock terms 'emotional reactivity,' and (b) a bipolar factor, distinguishing what may be loosely called the extravert type from the introvert type (the word 'type' designating a tendency of varying degree); (iii) if we ignore the more complex components acquired during the individual's life-time (interests, sentiments, ideals, etc.), the more important differences between one personality and another can be usefully assessed in terms of these two main components. (To these, however, I myself would add a third, namely, a second bipolar factor distinguishing the cheerful, euphoric type from the gloomy dysphoric). I believe that some such formulation would also be accepted by independent authorities, such as Allport, Guilford, and Vernon. The chief differences would, I believe, mainly be differences of wording or nomenclature. In particular, I should deprecate the marked preference shown by most psychological writers for a psychiatric terminology even when discussing normal individuals.

IV.—PRACTICAL COROLLARIES.

Let us now pass to the practical implications of these theoretical conclusions, and more particularly their relevance to the daily work of the teacher. First of all, there can be little question that the basic qualities of a child's temperament and character have a far-reaching influence on his educational progress while still in a *statu pupillari*. Current discussions about the large number of intelligent pupils who fail to reach a grammar school, or to enter a university, or to pass their degree

examinations, too often overlook this point. Studies of children who are educationally backward reveal positive correlations of both 'emotional instability' and 'emotional apathy' with a lack of industry and general progress, and of extraverted tendencies and introverted tendencies, with failure in certain specific subjects. Studies of failure among scholarship and gifted children produce much the same correlations.

But the teacher who is genuinely interested in the pupils in his charge is concerned, not merely with each child's intellectual progress at school or at college, but also with his development and prospects in later life. It has long been a commonplace thing with educational writers that (in the words of one of the most eminent—Sir Percy Nunn) "the education given in the elementary school, like that given in the public school, should aim at training the character of the pupils as well as instructing their intellects." During my period of work in the London schools, it was a deliberate policy of the Inspectorate to encourage teachers to take note, not only of their pupils' differences in intelligence and special aptitudes, but also of their varying qualities of temperament and character. This, so it was argued, "is the surest way to diminish the number of future criminals and delinquents and to forestall neurotic breakdown."

No psychological tests or techniques, no interview by a psychologist or a clinical psychologist, can yield such trustworthy assessments or such sound predictions as those of an experienced teacher who has watched and studied his pupils' behaviour and development, year after year. Experience, however, is essential. By far the best judgments are those of older headmasters and headmistresses who have kept in touch with their pupils after they have left school, either through local clubs or 'old pupils' associations,' or, best of all, actually residing in the same district as their pupils. But the majority of teachers have neither gained this prolonged experience nor enjoyed these opportunities; and to them the psychologist can offer a good deal of help. As Vernon (1953) has insisted, "the identification of what to measure is the first essential step towards a science of personality . . . With half-a-dozen or so major dimensions accurately measured, we could do a much better job than we can by naive intuition or clinical methods." Despite their obvious imperfections, factorial researches have done much to indicate what characteristics are most likely to be relatively permanent, what are the most useful points to observe, and how to record and assess them. "Such factors have the advantage of better definition and distinctiveness than the layman's usual categories" (Vernon, 1964).

The reliability and validity of the various procedures now available I have discussed in earlier issues of this *Journal* (1945 *a* and *b*). Cattell and Eysenck strongly favour 'objective tests'; but here I fully agree with Vernon's criticisms. The reliability of the best of them is well below 0.60, and their validity lower still. An interview conducted by an appropriately trained educational psychologist is far more effective. For the practical teacher, a questionnaire (to be filled in by the observer, not by the subject) will provide reasonably reliable results, at least for a preliminary screening. And these can then be supplemented by more elaborate modes of observation along the lines I have elsewhere described.

To illustrate the value and the possibilities of such methods, may I briefly report the results of a follow-up carried out in two London areas? Here I shall confine myself chiefly to predictions of delinquency and neurosis: these are failings which are most directly related to temperament and character, and lend themselves to fairly definite statement and calculation. It so happens that I have lived on the edge of one of the districts concerned for over forty years, and for much of that time was an active member of a Settlement in the other area. Thus, my co-workers and I became closely acquainted with the families who remained in those areas during the period in question. The follow-up procedure was seriously interrupted during the last war; but, when the war was over, we were able to trace a large proportion of those in whom we had been chiefly interested even though they had moved elsewhere. The data here recorded relate solely to boys: to keep in touch with the girls proved far more difficult and far less informative. With most of the children

the post-school period of observation extended over at least six years; with some of the earliest cases it has lasted for more than thirty years. The total number of children for whom we obtained reasonably reliable assessments amounted to 763. Of these the vast majority, so far as could be ascertained, followed perfectly normal and law-abiding careers.

For the initial assessments we were indebted chiefly to the observations and ratings of teachers, four of whom had taken university courses in psychology, and all of whom had attended L.C.C. lectures on "The Assessment of Ability and Character." I interviewed every child whose assessments deviated appreciably from the general average, and used my own assessments as a means of equating the standards of different teachers. A child was classified as 'highly emotional' if the assessment for 'general emotionality' placed him among the 25 per cent. most emotional children in the group. A similar borderline was adopted for the 'unstable.' Just under half of the 'highly motional' children were classified as 'extraverted' or 'unrepressed'; and a somewhat smaller proportion as 'introverted' or 'repressed': (see Table 1).

TABLE 1

A COMPARISON OF THE PERSONALITY ASSESSMENTS OF SCHOOL CHILDREN WITH THEIR AFTER HISTORIES.

The numbers are expressed throughout in percentages.

School Assessment	Habitual Offenders		Other Delinquents		Neurotics	
	Assessed among Actual Cases	Actual Cases among Assessed	Assessed among Actual Cases	Actual Cases among Assessed	Assessed among Actual Cases	Actual Cases among Assessed
High Emotionality	63	15	51	19	59	18
Emotional Instability	67	16	56	21	53	16
Unrepressed (Extraverted) ..	54	27	38	30	14	9
Repressed (Introverted) ..	3	2	14	11	44	29
Final Predictions	49	83	36	67	31	58

The teachers were also to pick out those whom they thought would develop into (a) habitual offenders, (b) possible or occasional offenders, or (c) cases of neurotic breakdown. Of the first group 83 per cent., roughly five out of six, actually did drift into a life of regular and persistent crime. Nevertheless, half of the habitual cases had not been recognized as such by the teachers, though some were included among the 'possible' offenders and many of the remainder had been assessed as 'unstable'—usually of a so-called 'extraverted' type. In predicting the 'occasional' cases, the teachers were far less successful, no doubt because the commission of a single offence depends so much on the mood and circumstances of the moment. Of those who eventually suffered some kind of nervous breakdown, 59 per cent. (well over twice the proportion found among the total population) had been assessed as highly emotional during their school years. And of these who had been assessed as 'highly emotional,' 18 per cent. became neurotic or psychotic; but four-fifths showed no subsequent signs of any neurotic tendency. Rough as they are, the figures clearly demonstrate that an early attempt at noting the more emotional and unstable individuals does provide some indication as to which pupils need to be watched and perhaps recommended for special study or treatment. At the same time, it implies the additional influence of various unforeseen precipitating factors, arising at a later stage.

In addition to these studies of delinquency and neurosis, we were able to obtain full and accurate data regarding the occupation, earnings, and general progress of the vast majority of our cases. Here, too, it became clear that the assessment of the character-qualities, carried out while the children were still at school, did much to throw light on the causes of failure or success. But it would be difficult to reduce these further data to a succinct or convincing summary. I will merely say that, in my view, owing to a constant pre-occupation with the more sensational problems of crime and neurosis, the psychologist's study of personality has in the past been too exclusively concentrated on abnormal aspects, and has tended to neglect the influence of affective and conative factors on the cultural, social, and vocational progress of the average child. Moreover, we have been too easily satisfied with short-term inquiries and with surveys based simply on cross-sections of the school population. What are now required are series of longitudinal studies, systematically planned on a co-operative basis with an eye to following-up a large and representative sample of the school population year by year, well into adult life. In these researches, particularly at the initial stages, the interest and help of the school teacher will be indispensable. For the later stages no doubt we shall need to invoke the aid of the social worker, the social psychologist, and the industrial psychologist, continuing the observations under the guidance of some permanent organization for psychological research.

V.—REFERENCES.

- ADCOCK, C. J. (1965). A comparison of the concepts of Cattell and Eysenck. *Brit. J. Educ. Psychol.*, 35, 90-97.
- ALLPORT, G. W., et al. (1946). Personality: a symposium. *Brit. J. Educ. Psychol.*, 16, 57-68, 105-116.
- BANKS, C. (1948). Primary personality factors in women: a re-analysis. *Brit. J. Psychol., Stat. Sect.*, 1, 204-218.
- BURT, C., and MOORE, R. C. (1912). Mental differences between the sexes: (ii) emotional. *J. Exp. Pedag.*, 1, 355-388.
- BURT, C. (1915). The general and specific factors underlying the primary emotions. *Brit. Ass. Ann. Rep.*, 84, 694-696.
- BURT, C. (1917). The unstable child. *Child Study*, 10, 61-78.
- BURT, C. (1935). *The Subnormal Mind*. London: Oxford University Press.
- BURT, C. (1938). The analysis of temperament. *Brit. J. Med. Psychol.*, 17, 158-188.
- BURT, C. (1939). A factorial analysis of emotional traits. *Character and Personality*, 7, 285-299.
- BURT, C. (1945). The reliability of teachers' assessments of their pupils. *Brit. J. Educ. Psychol.*, 15, 80-92.
- BURT, C. (1945). The assessment of personality. *Brit. J. Educ. Psychol.*, 15, 107-121.
- BURT, C. (1946). The assessment of personality. *J. Egypt. Psychol.*, 2, 65-85.
- BURT, C. (1948). The description and measurement of personality. *Brit. J. Psychol., Stat. Sect.*, 1, 134-136.
- BURT, C. (1948). The factorial study of temperamental traits. *Brit. J. Psychol., Stat. Sect.*, 1, 178-203.
- BURT, C. (1949). The structure of the mind. *Brit. J. Educ. Psychol.*, 19, 100-111, 176-199.
- CATTELL, R. B. (1946). *The Description and Measurement of Personality*. Yonkers: World Book Co.
- CATTELL, R. B. (1948). The primary personality factors in women. *Brit. J. Psychol., Stat. Sect.*, 1, 114-130.
- CATTELL, R. B. (1957). *Personality and Motivation: Structure and Measurement*. London: Harrap.
- EYSENCK, H. J. (1947). *Dimensions of Personality*. London: Kegan Paul.
- EYSENCK, H. J. (1952). *The Scientific Study of Personality*. London: Kegan Paul.
- EYSENCK, H. J. (1957). *The Structure of Human Personality*. London: Methuen.
- EYSENCK, H. J. (1964). Principles and methods of personality description, classification, and diagnosis. *Brit. J. Psychol.*, 55, 284-294.
- JUNG, C. G. (1923). *Psychological Types*. London: Kegan Paul.
- KRETSCHEMER, E. (1925). *Physique and Character*. New York: Liveright.
- PAVLOV, I. P. (1941). *Lectures on Conditioned Reflexes*. New York: McGraw-Hill.
- REYMERT, M. L. (Ed.) (1950). *Feelings and Emotions*. New York: McGraw-Hill.
- SHIELDS, J. (1954). Personality differences and neurotic traits in normal twin school children. *Eugen. Rev.*, 45, 213-246.

- SHIELDS, J. (1962). *Monozygotic Twins*. London: Oxford University Press.
 TERMAN, L. M., and ODEN, M. H. (1947). *The Gifted Child Grows Up*. Stanford: Stanford University Press.
 VALENTINE, C. W. (1940). The specific nature of temperament traits and a suggested report form. *Brit. J. Educ. Psychol.*, 10, 25-48.
 VERNON, P. E. (1953). *Personality Tests and Assessments*. London: Methuen.
 VERNON, P. E. (1964). *Personality Assessment: A Critical Survey*. London: Methuen.
 WUNDT, W. (1910). *Grundzüge der Physiologischen Psychologie*. Vol. II. Leipzig: Engelmann.

(Manuscript received 22nd July, 1965).

PERFORMANCE OF CANADIAN SCHOOL CHILDREN ON THE J.M.P.I.

By

A. J. CROPLEY

(University of New England, Australia)

AND

B. R. D'AOUST

(University of Alberta, Canada)

I.—SUBJECTS.

Subjects were 320 Canadian children who comprised the entire Grade Seven class of a large metropolitan junior high school in the city of Edmonton, Alberta, Canada. The sample included 170 boys, whose mean age was 13 years 10 months (range 11 years 1 month to 16 years 1 month; SD=16.6 months), and 150 girls with a mean age of 13 years 2 months (range—11 years 7 months to 16 years 1 month; SD=17.3 months). Mean I.Q. of the boys was 114.7 (range —79 to 148; SD=15.2), while the mean I.Q. of the girls was 113.8 (range —88 to 150; SD=13.6). Corresponding figures for the full sample of 320 were 13 years 6 months (range —11 years 1 month to 16 years 1 month; SD=16.9 months) and 114.3 (range —79 to 150; S.D.=14.5).

II.—RESULTS.

The present author's findings are reported in Table 1, and compared with those of Furneaux and Gibson (1961).

TABLE 1
SCORES OF ENGLISH AND CANADIAN CHILDREN ON THE J.M.P.I.

	n	E		N		r
		Mean	SD	Mean	SD	
Furneaux and Gibson	156	12.39	3.46	7.35	3.54	— .10
Cropley and D'Aoust	320	12.05	3.25	7.57	2.94	— .085
Cropley and D'Aoust:						
Boys	170	12.16	3.13	7.44	2.99	— .067
Girls	150	11.92	3.38	7.71	2.89	— .095

The outstanding characteristic of these figures is the remarkable similarity between means of both English and Canadian samples. None of the Canadian means reported here differs significantly from the corresponding Furneaux and Gibson score. Furthermore, the correlations between E and N failed even to approach levels of statistical significance (full sample: $r = -.085$; boys: $r = -.067$; girls $r = -.095$).

III.—DISCUSSION.

The present data fail to support the suggestion that Canadian children differ from English in their mean scores on the JMPI, which has been advanced by Costello and Brachman (1962). Furthermore, while the correlations between E and N were negative as they have been in a number of other studies (Furneaux and Gibson, 1961; Gallard and Goodfellow, 1962; Costello and Brachman, 1962), they are of virtually zero magnitude, and strongly support the notion that the two scales are orthogonal. Again, this finding may be contrasted with that of Costello and Brachman, who reported (1962) a substantial correlation between the two subscales ($-.179$), and questioned the independence of the E and N dimensions.

Consequently, the present findings are strongly supportive of the use of the JMPI with appropriate groups of Canadian school children, suggesting that it is an equivalent measure in both Canada and England. Although it is true that mere group means tell nothing about the way in which scores were obtained, the highly comparable means of the Canadian and British samples suggest that research involving the JMPI in both Britain and North America probably yields results which are directly comparable.

IV.—REFERENCES.

- CALLARD, M. P., and GOODFELLOW, C. L. (1962). Neuroticism and extraversion in schoolboys as measured by the J.M.P.I. *Brit. J. Educ. Psychol.*, 32, 241-250.
- COSTELLO, C. G., and BRACHMAN, H. M. (1962). Cultural and sex differences in extraversion and neuroticism reflected in responses to a children's personality inventory. *Brit. J. Educ. Psychol.*, 32, 254-257.
- FURNEAUX, W. D., and GIBSON, H. B. (1961). A children's personality inventory, designed to measure neuroticism and extraversion. *Brit. J. Educ. Psychol.*, 31, 204-207.
- GIBSON, H. B. (1964). A lie scale for the Junior Maudsley Personality Inventory. *Brit. J. Educ. Psychol.*, 34, 120-124.

(Manuscript received 31st May, 1965.)

I.Q. CHANGES IN EDUCATIONALLY SUBNORMAL CHILDREN AT SPECIAL SCHOOL

By

ZENA A. STEIN*

(Research Fellow, Department of Social and Preventive Medicine, University of Manchester)

AND

GREGORY STORES†

(Medical Student, University of Manchester)

There is increasing evidence that among those people classed as subnormal in childhood, certain groups show intellectual improvement as adults. Clarke and Clarke (1953, 1954), and Clarke, Clarke and Reiman (1958) show that I.Q. increments frequently occur in feeble-minded patients who have suffered extreme deprivation

* Zena Stein holds a personal grant from the Medical Research Council.

† Gregory Stores is aided by a grant from the Sir Halley Stewart Trust.

in early life. Mundy (1957) presents similar findings. Stein and Susser (1960) found that clinically normal subjects, reassessed as young adults, made I.Q. increments whereas subjects with definite or presumed neurological lesions did not.

In the present investigation we were interested to know if I.Q. changes had occurred by the time E.S.N. children left school and what the predictive value was of intelligence test results obtained when the children were ascertained as educationally subnormal at school. The present evidence is conflicting (Scarr, 1953; Collman and Newlyn, 1958; Hiskey and Sadnavitch, 1958; Rushton and Stockwin, 1963). I.Q.s were obtained before admission and again before leaving for a mixed group of seventy-five children attending a special day school for E.S.N. children in Lancashire. The average age on initial testing was 11.2 years; that on final testing was 16.0 years. The comparisons are shown below.

TABLE 1

	Initial Testing : Terman-Merrill I.Q. (with Roberts-Mellone Correction) (Roberts and Mellone, 1952)			Final Testing : Wechsler Full Scale I.Q.		
	Boys (n=45)	Girls (n=30)	All (n=75)	Boys	Girls	All
Mean	74.3	71.2	73.1	76.2	71.9	74.5
S.D.	6.38	8.17	7.22	11.73	11.79	11.76
Range	59-85	57-87	57-87	50-102	55-99	50-102
Av. final I.Q. : minus av. initial I.Q.:						
Boys ..	+1.9 Difference not significant.					
Girls ..	-0.7 Difference not significant.					
All	+1.3 Difference not significant.					
Difference between boys' av. initial I.Q. and girls' av. initial I.Q.	3.1 Not significant.					
Difference between boys' av. final I.Q. and girls' av. final I.Q.	4.3 Not significant.					
Test—Retest Interval :						
Mean ..	4.5 years.					
S.D. ..	1.8 years.					
Range ..	0.5—9.8 years.					
Product Moment correlation between Terman-Merrill (with Roberts-Mellone correction) and Wechsler Full Scale scores = +0.71. Level of significance : $P < 0.001$.						

When Terman-Merrill and Wechsler I.Q.s. of mentally subnormal children of the age range in question are compared, the Wechsler Full Scale score is, on average, higher than the Terman-Merrill score by about two points (see review by Rohrs and Haworth, 1962). After correcting for this difference, the magnitude of change between initial and final I.Q. was examined for each child. If we take a difference of ± 3 points, after this correction, as a reasonable measure of retest error, three groups can be distinguished: Improved, with I.Q. gains of between 4 and 22 points ($n=24$); Unchanged, with test-retest differences of ± 3 points ($n=24$), and Deteriorated, with I.Q. decrements of 4 to 18 points ($n=27$).

These I.Q. discrepancies suggest that the predictive value of the Terman-Merrill score is limited and that a constancy of I.Q. is seen in only one-third of our sample whereas, by the time school leaving age has been reached, marked changes have occurred in the rest—in a few cases sufficient to lift the children well into the range of intellectual normality.

REFERENCES.

- CLARKE, A. D. B., and CLARKE, A. M. (1953). *Lancet*, 877-880.
 CLARKE, A. D. B., and CLARKE, A. M. (1954). *Brit. J. Psychol.*, 45, 173-179.
 CLARKE, A. D. B., CLARKE, A. M., and REIMAN, S. (1958). *Brit. J. Psychol.*, 49, 144-157.
 COLLMAN, R. D., and NEWLYN, D. (1958). *Am. J. Ment. Defic.*, 63, 2, 307-311.
 HISKEY, M. S., and SADNAVITCH, J. M. (1958). *Exceptional Children*, 25, 16-20.
 MUNDY, L. (1957). *Brit. J. Med. Psychol.*, 30, 194-201.
 ROBERTS, J. A. F., and MELLONE, M. A. (1952). *Brit. J. Psychol. (Stat. Sect.)*, V, ii, 65-79.
 ROHRS, F. W., and HAWORTH, M. R. (1962). *Am. J. Ment. Defic.*, 66, 853-859.
 RUSHTON, C. S., and STOCKWIN, A. E. (1963). *Brit. J. Educ. Psychol.*, 33, 132-142.
 SCARR, E. H. (1953). *Brit. J. Psychol. (Stat. Sect.)*, 6, 71-76.
 STEIN, Z. A., and SUSSER, M. W. (1960). *J. Ment. Sci.*, 106, 1296-1319.

(Manuscript received 26th April, 1965.)

EXTRAVERSION, NEUROTICISM AND THE EFFECT OF PRAISE OR BLAME

By Y. RIM

(Technion—Israel Institute of Technology*)

In their excellent review of thirty-five articles spanning fifty years of research on the effects of praise and blame on the performance of school children, Kennedy and Willcott (1964) point out that the first study, which attempted to take into account personality differences was that by Forlano and Axelrod (1937). Their investigation at the Fifth Grade level has indicated that repeated applications of praise or blame have differential effects on the work performance of introverts and extraverts. They found blamed introverts significantly superior to praised or control introverts. They repeated praise or blame only twice, and concluded that blame is more effective than praise, but speculated as to whether more trials would have brought about significance for the praise groups. The second study on the effects of praise and blame on the work achievement of introverts and extraverts at the same grade level was that of Thompson and Hunnicutt (1944). Whereas Forlano and Axelrod used three trials, Thompson and Hunnicutt used six repeated measures under the conditions of praise and blame. Thompson and Hunnicutt concluded that when introverts and extraverts are grouped together, praise and blame are equally effective in motivating the work achievement of their subjects, as compared with no external incentives. In addition, Thompson and Hunnicutt found that if repeated often enough, praise increases the work output of introverts until it is significantly higher than that of introverts who are blamed or extraverts who are praised.

* This investigation was carried out while the author was spending his Sabbatical leave at the Department of Psychology of the Institute of Psychiatry, University of London. Thanks are due to Professor H. J. Eysenck and Dr. S. B. G. Eysenck for their encouragement and helpful discussions; to Mr. M. Berger for his assistance in testing; and to Mr. Potter, Headmaster of Forest Hill School, and Miss Love, Headmistress of Sydenham Secondary School, for their permission to carry out the testing at their schools.

This seemed to answer in the affirmative Forlano and Axelrod's question as to whether repeated trials would raise the effects of praise to significance. Repeated blame increased the performance of the extravert-blame group until it was superior to the extravert-praise and introvert-blame group.

Thompson and Hunnicutt comment rightly: the results of this study indicate that praise, as well as blame, can be used unwisely by the elementary school teacher if he does not fully appreciate and understand the different personalities present in his classroom. Praise and blame should not be judged on an either/or basis, but should be used to fit the case.

The basic work done by Cattell (1961), Eysenck (1960) and Guilford (1959) has shown that there is strong support for a two-dimensional model of personality: Extraversion-Introversion (E), and Neuroticism, emotionality or stability-instability (N). The present study's aim is twofold: (a) to replicate the two above-mentioned studies by using the Extraversion scale of the Junior Eysenck Personality Inventory, and (b) to find whether the use of the Neuroticism scale from the same inventory would add to our understanding of the effect of praise or blame on children's performance.

METHOD.

Subjects. Twelve-year-old pupils (sixty-eight girls and seventy-five boys) from two comprehensive schools served as subjects. In each school three forms were selected: two served as experimental groups and one as control group. The forms were of similar ability levels.

Tests used. The Junior Eysenck Personality Inventory (1965) was administered to all subjects and scored for Extraversion and Neuroticism. The J.E.P.I., designed for use with children from the ages of 7 to 16, was developed by extensive factor analyses carried out separately for the various age groups, and detailed norms are available for boys and girls separately. Two scores were derived for each child: one for Extraversion (E) and one for Neuroticism or emotionality (N).

Task and Procedure. Six alternate forms of a cancellation test were used to measure each pupil's work achievement under the various experimental conditions. Each of the cancellation tests consisted of randomly distributed capital letters of the alphabet. The children were instructed to cancel each capital (E). The duration of each test was thirty seconds. In the Control Group no comments were made between any of the six testing sessions. In the experimental groups, however, the experimenter passed around the room after each testing period and said, after having studied each pupil's paper for a second or two: "Very well," "Good," "Well done," "Excellent," etc., in the groups to be praised. In the groups to be blamed, he said: "Not very good," "Still not good enough," "This is definitely not good enough," etc. In addition to verbal expressions of disapproval or approval, accepted social gestures, such as nodding or shaking of the head, were used.

Results and discussion. Preliminary analyses showed that there were no significant differences between the data of the boys and girls; therefore, the data were combined and the results will be presented for both groups.

In general, our results tend to confirm that praised or blamed children improve their performance, in comparison with a control group. But it is obvious from Table 1 that blamed children, improve their performance significantly more than praised children, and that not all praised children improve their performance. Thus, children of high 'E' and low 'N' scores tend to improve more than controls, children of low 'E' scores improved less and those of high 'N' scores show no appreciable difference.

It should be mentioned that none of the differences between the praised and control groups reach statistical significance, whereas all the differences between the praised and blamed groups are significantly different at the .05 level or better. Similarly, all the differences between the blamed and control groups are significantly different at or beyond the .05 level.

TABLE 1
MEAN PERFORMANCE OF PRAISED AND BLAMED CHILDREN, ACCORDING TO EXTRAVERSION
AND NEUROTICISM.

	CONTROL					PRAISED					BLAMED				
	'n'	1st 3 trials	Last 3 trials	Diff.	't'	'n'	1st 3 trials	Last 3 trials	Diff.	't'	'n'	1st 3 trials	Last 3 trails	Diff.	't'
High 'E'	14	17.6	17.5	(-0.1)	1.61	26	19.7	22.3	(+2.6)	1.62	25	19.0	28.0	(+9.0)	*2.55
Low 'E'	10	15.9	18.0	(+2.1)	1.72	23	19.6	19.9	(+0.3)	1.41	15	19.6	24.9	(+5.3)	1.72
High 'N'	10	18.4	19.0	(+0.6)	1.58	21	18.7	19.5	(+0.8)	1.50	16	20.7	26.0	(+5.3)	1.81
Low 'N'	14	15.4	17.1	(+1.6)	1.61	28	19.0	23.0	(+4.0)	*2.34	24	18.6	27.2	(+8.1)	*2.42

* 't'—Significant at the .05 level.

It is possible that our way of praising was too mild, or too similar to the way of day-to-day encouragement at school, whereas the blaming was unusual, and, therefore, a strong incentive to improve performance.

Table 1 shows that three significant differences were found between the performance on the first three trials and that of the last three trials: for children of low

'N' scores (below the median of their forms), whether praised or blamed, and for children of high 'E' scores (above the median of their forms), who were blamed. All these children significantly improved their performance in the last three trials, as compared with the first three trials.*

This result is not entirely in accordance with the results achieved by Thompson and Hunnicutt, who found that praised introverts as well as blamed extraverts would do better than blamed introverts and praised extraverts. It is true that we, too, found blamed extraverts improving their performance more than blamed introverts, but our results do not support the finding that praised introverts do better than blamed introverts. If anything, our results suggest that praised extraverts do better than praised introverts, although the difference fails to reach an acceptable level of significance.

It seems that the additional information showing that children of low 'N' scores improve their performance, whether blamed or praised, helps us to interpret Thompson and Hunnicutt's findings.

Children of high 'E' and low 'N' scores tend to improve their performance when blamed, whereas praised children tend to improve their performance if their 'N' score is low independently of whether they are introvert or extravert. Thompson and Hunnicutt's findings could be due to their use of a test of introversion/extraversion which might in part at least have been measuring neuroticism. Unlike the Junior E.P.I., it was not based on extensive item and factor analyses.

We agree with Thompson and Hunnicutt that praise and blame should be used to fit the case, and that the teacher may use them unwisely if he does not fully appreciate and understand the different personalities in his classroom. It seems to us that knowledge of a child's degree of emotional stability as well as his degree of extraversion/introversion can improve a teacher's judgment in which cases to use praise or blame.

SUMMARY.

The influence of praise and blame on 12-year-old children was studied. Previous investigations found that praised introverts and blamed extraverts improve their performance more than blamed introverts and praised extraverts. The present study, using measures of both extraversion/introversion (E) and neuroticism or emotional stability-instability (N), shows that blamed extraverted and emotionally stable children, whether blamed or praised, improve their performance. The importance of using both personality dimensions of 'E' and 'N' for making decisions in the use of praise and blame is stressed.

REFERENCES.

- CATTELL, R. B., and SCHEIER, I. H. (1961). *The Meaning and Measurement of Neuroticism and Anxiety*. New York: Ronald.
- EYSENCK, H. J. (1960). *The Structure of Human Personality*. London: Methuen.
- EYSENCK, S. B. G. (1965). A new scale of personality measurement in children. *Brit. J. Educ. Psychol.*
- EYSENCK, S. B. G. (1965a). *The Junior Eysenck Personality Inventory*. London: Univ. London Press.
- FORLANO, H., and AXELROD, C. (1937). The effect of repeated praise or blame on the performance of introverts and extraverts. *J. Educ. Psychol.*, 28, 92-100.
- GUILFORD J. P. (1959). *Personality*. New York: McGraw-Hill.
- KENNEDY W A., and WILCOTT H. C. (1964). Praise and blame as incentives. *Psychol. Bull.*, 62, 323-332.
- THOMPSON, G. G., and HUNNICUTT, C. W. (1944). Effects of repeated praise or blame on the work achievements of introverts and extraverts. *J. Educ. Psychol.*, 35, 257-266.

(Manuscript received 22nd June, 1965.)

* Throughout this study the significance of the differences between means were tested by t-tests for independent samples or correlated samples, as appropriate.

CRITICAL NOTICE

UNIVERSITY GRANTS COMMITTEE (1964). *Report of the Committee on University Teaching Methods*. London: Her Majesty's Stationery Office, pp. 173, 10s.

By T. H. PEAR

This valuable, moderately priced publication differs from many official reports in that it will interest many readers who are not professionally concerned with its theme. Using both subjective and objective (statistical) evidence, it illuminates important subjects which for years have been hotly discussed in many a common room, not only in this country. Many university members and many school teachers whose syllabuses are directed or corseted by examiners 'up there' might usefully ask themselves "What questions would I expect this committee to consider? How many are old but unsolved, how many arise out of present situations; how many concern the foreseeable future?"

Though not pretending that my own answers are typical, I suggest that few are idiosyncratic. They follow extracts from the Report, each with *ad hoc* comments or criticisms. As the committee contained three vice-chancellors and a head mistress, problems of administration are not neglected. There is an excellent summary, from which the present reviewer has quoted sentences, usually literally:

"We regard as the main purpose of the first-degree course . . . to train the student to think for himself and work on his own . . . the vacations should be of value in giving him practice (in these activities)."

The Committee did not examine a current belief, not universally shared, that a considerable part of the vacation might properly be spent in earning money to supplement grants or to pay for a holiday abroad, with possible educational benefits. In an interim report (1963):

"We discussed the difficulties, financial and other, which students found in using the vacation to best advantage, and made a number of suggestions which we believe are being followed up."

Certain features which influence the teaching and supervision of students are discussed. One is:

"The prestige of specialised honours courses and the loss of esteem, particularly in England, of more generalised courses." (There is criticism of) "the practice of some universities of admitting students direct to honours courses and later relegating to general courses students of marginal performance . . . The alternative, universal in Scotland and Wales, of admitting students to general courses and promoting some to specialised work for honours at a later stage, involves problems of its own . . . The structure of the ordinary degree course in Scotland necessarily involves very large classes . . . With the change recently made in the regulations governing university awards, more flexibility in permitting changes of subjects might be possible."

Psychological reasons behind a student's change of course after a few terms at the university could be more carefully studied: a few sudden conversions happen, by no means always in students who are doing only moderately well; for example, some first-year honours students 'good at' one of the physical sciences may find that the ways in which knowledge is attained and used now interest them more than isolated facts. Study of language may lead to study of psychology, anthropology or communication.

"In some degree courses the extent of a student's success in sessional examinations held in June determines what work he will do in the next session. This may be unknown to him when he leaves for the long vacation, and may prejudice his use of it. Moreover, our inquiries into students' use of the vacations showed that students expecting to sit examinations or undergo other tests in the Autumn are more diligent in the long vacation than other students. For such reasons there would be advantage to teaching in making the academic year run from January to December."

Such a change has also been advocated on other grounds. It would involve complex administrative issues which are being separately examined. We think that this problem in all its aspects should be considered as a whole."

Since my comments are written for a psychological journal, some can be made from the standpoint of a social psychologist. As one who spent his last undergraduate's long vacation happily in reading books and articles often marginally connected with psychology, and in the final examination gleefully attempted an essay on a subject which I had explored for fun, I naturally agree with the above suggestion. But there are students whose parents and neighbours regard "all this reading", as a waste of time, and a convenient dodge for avoiding some real work. There is too, the more serious claim that much of the non-academic work done in the vacation trains the character. Probably, but sometimes in queer ways. One of our toughest writers—ex-service and Cambridge—doubted whether the student he saw taking generous tips from casual passengers in the lift of a luxury hotel was training his character desirably. And many years ago the long vacation was regarded by university teachers as the period when the student could read the books and journals recommended to him in lectures.

"In most universities the basic units of organisation are the departments, which enjoy a large measure of independence. This form of organisation involves certain problems . . . Perhaps the most serious is the difficulty of providing adequate supervision for the student who belongs wholly to no department and thus may feel himself to be a 'second-class citizen.' If more students are to take broader courses, as recommended by the Committee on Higher Education, these problems will become increasingly important. We think that in some universities the incompatibility between full departmental autonomy and the satisfactory conduct of multi-subject courses requires further study."

To the director of a department this independence can be very enjoyable, it offers other, less personal advantages, but it allows him, if he wishes, to obstruct, covertly and overtly, the introduction into the university curricula of a new subject—new in his university though middle-aged in some others, by claiming that 'it' is provided for in his and some other departments. The year-long successful resistance to Sociology is a case in point.

"The transition from school to university requires the student to make a rapid adjustment in his way of life and in his approach to learning, and his first year at the university is, therefore, a critical period for him." (Some account is given of) "the extent to which students are prepared for this change before and at the beginning of their university life, and during their first year at the university."

It is not surprising that 'student opinion' as expressed in reports from students' organisations is not always consonant with statistical evidence, since students' organisations contain not only members whose outlook is unusually thoughtful and reformist, but also temperamental revolutionaries.

"Students' opinion is critical both of the lecture as a method of teaching, on the grounds that it is a 'one-way' process, and also of the quality of some lectures. The value of some types of lecture is recognised, but the student organisations think that too much emphasis is placed on lectures. The statistical evidence, however, suggests that the majority of individual students do not wish for a change in the time given to lectures. Of those who wish for change, most want fewer but there is an appreciable minority who want more. All the evidence is at one in showing a strong student demand for more teaching by tutorial and seminar. There is relatively little comment on other methods of teaching."

Neither in the Report nor in recent newspaper articles is it suggested that a particular tutor's personal qualities, his degree of up-to-dateness, his freedom from conventional prejudice and his motives for spending a great deal of time in tutorial guidance to individual students are of psychological interest. We read occasionally that the student and the tutor to whom the student has been assigned may have incompatible temperaments, and hear that occasionally a student is dissatisfied with the tutor's knowledge of fringe-subjects on which he gives advice. If a critical student for the last few years has habitually listened to specialists' talks on the

radio and television, may he not at times be ahead of a hard-working tutor? What, if anything, does he do about it? Not only questions of tact may be involved here.

"... the overwhelming weight of university (teacher's) opinion is that lectures have an essential function, particularly for opening up a subject for students who are not in a position to do it for themselves by unassisted reading, and also for giving more detailed information where suitable textbooks are lacking. Lectures have certain advantages over discussion periods in that a continuous exposition, free from interruptions, can be better prepared and more profound than teaching in a discussion period, can cover more ground, and can enable an inspiring teacher to influence more students." ... "We take the view that students should be clearly informed whether their attendance at a particular set of lectures is obligatory."

On this last question, the report from the National Union of Students of England, Wales and Northern Ireland (1961) says unequivocally:

"It has been brought to our attention that at a great number of universities and university colleges, attendance at lectures is compulsory, if only nominally in some cases. We believe this to be a most unfortunate infringement of the student's academic freedom, believing that a student should feel himself as equally free to choose which lectures he goes to as he is to select which books he reads..."

Is this assertion seriously meant to apply to students of, say, medicine or engineering?

In considering methods of teaching other than the lecture, the Report valuably compares seminars with tutorials:

"The former tend to be subject-centred, and the latter to be student-centred, but both commonly begin with the presentation of a paper or essay on a set topic followed by discussion. It is widely held that numbers at a tutorial should not exceed four, and many prefer individual teaching. For seminars the size of group preferred ranges between six and twelve, but many seminars are larger. At Oxford and Cambridge tutorials are smaller than elsewhere, but arts students at those universities do not spend more time in discussion periods attended by nine or fewer students than arts students elsewhere. In most sciences, on the other hand, tutorial teaching is relatively little used outside Oxford and Cambridge."

Since popular discussion of the 'tutorial' often centres on or is confined to a system traditional at Oxford, an account of it is given:

"... as representing one end of the spectrum. There, nine out of ten tutorials are given to one or two students, and responsibility for organising the student's work and giving some teaching is normally concentrated in one tutor. These features have both advantages and disadvantages which we discuss... The case for the discussion period is that, rightly handled, it provides for active participation by the student and for immediate assessment of his efforts... The memoranda which the Bridges Syndicate at Cambridge received from senior tutors, agreed that supervision 'is directed to fostering an undergraduate's general intellectual development, by stimulating his interests and widening his outlook, as well as by overseeing his general progress and, perhaps above all, by teaching him how to make the best of his abilities and to learn how to study and master a subject for himself.' However, there is a tendency for supervision to assume much more the character of a small instructional class, of which the purpose is to make certain that the undergraduate has properly understood and learned the facts in a particular branch of his subject."

In other words, coaching. The syndicate speak of the need to save supervision "from possible deterioration into coaching for the Tripos," and favour a reduction in the amount of supervision. The Report comments:

"Some universities appear to use discussion periods in a way which is not easy to distinguish from the 'coaching' of which the Bridges Syndicate speak with disapproval... The view that the main purpose of small-group teaching is to remove student's difficulties is implicit in the comment made by one university that it is 'unsuitable for the very clever student who has few difficulties.' There seem to us to be two respects in which coaching may be open to objection... a form of teaching which is too closely geared to examinations, for example, where the student is taught more with a view to the supposed probability of particular questions being

set than to give him a connected grasp of his subject as a whole . . . The avoidance of such teaching is to some extent a problem for examiners . . . The term 'coaching,' when used in a pejorative sense, evokes a picture of a student who looks to his coach rather than to his own efforts to maintain his momentum and to help him over the difficulties of his course."

On examinations :

" . . . their scope and character, and the questions which students must answer, should be such that in preparing for examinations they will not be able to avoid being educated. This may make it appear that the examiner is the dominant partner, but in our universities, teachers and examiners are, in general, the same people."

The functions of external examiners, if any, do not appear to have been discussed.

On the importance of university teaching, as a matter for training and study :

" . . . nearly three in five of our sample thought that newly appointed university teachers should receive some form of organised instruction or guidance on how to teach. We do not favour any prolonged course of training, but we think present arrangements unduly haphazard . . . there is, in our view, a good deal more that might be done to help the inexperienced in the skills of giving tutorials and lectures. Such help should be given as a matter of course to all newly appointed staff, without waiting until complaints are heard or mistakes have hardened into habits . . . Relatively little experiment on methods of university teaching has so far been carried out in this country. This we believe to be because there is little to tempt a university teacher to give to a study of teaching methods time which inclination and self-interest would lead him to give to research in his own subject. In America large sums have been made available by the Ford Foundation for operational research in teaching, including college and university training . . . "

A perennial subject for discussion at different levels of the university hierarchy is the emphasis placed by different university employers on teaching and research, respectively. When I came to Manchester to develop a department of psychology, a junior lecturer offered me an orientation talk. He said : " Of course, you'll be expected to 'produce'." " What ?," I asked. The dark answer was " You'll find out." I did, but since my last assignment had been to Oswald Külpe's laboratory at Würzburg, I probably assumed that both teaching and research were expected in all universities. I was therefore surprised when a colleague as young as myself intervened in a philosophical discussion with " A.B. always said . . . " To a poker faced man's rhetorical question : " Who's A.B. ? What's he written ? " came a calm reply : " He doesn't write ; he tutors."

While one heard from ' natural ' scientists few analyses of the value of research-in-general, there was even then a tendency among some Arts people to wonder how the advancement of knowledge could best be handed on to others. Scientists, at least the younger ones, tended to regard their laboratories (not yet called ' ivory ') as eyries, sanctums, even sanctuaries. An iconoclastic, hemlock-defying American had not then written *Knowledge for What ?* Knowledge ' must ' be advanced—one of those risky metaphors which visualisers often distrust—for ' its ' own sake. The cool use of ' it ' suggested that deification might be succeeding reification, though to some leading scientists metaphysics was ' just hot air.' ' Pure ' knowledge—a heavily loaded adjective—was still distinguished from applied, and in this connection mention of Leonardo da Vinci was in poor taste. The toast " And may this knowledge never be of use to anyone " was widely quoted by simple souls.

Even then, some debaters with feet on the ground maintained that research and publication were the only steps up the academic ladder. Few seemed surprised that candidates, admitting that they had never lectured were appointed with title of lecturer and clearly told to learn on the job. Ribald comparisons with air-pilots, dentists and violinists were commonly made. Some departments appointed as a lecturer, a person who was expected to give as few lectures as seemed prudent, but " What's in a name " ? : May week quite properly comes in June.

"Teaching and research must always go hand-in-hand," was often heard: but always in the same person? Was this assertion an axiom or a postulate, uttered so emphatically and so often that it was seldom examined? But what about Plato, Socrates, the dedicated tutors in Balliol's golden days? The report touches only the surface of such questions. Opinions of successful producers of radio-programmes, and especially on scientific subjects, would be valuable in this connection.

There was the evergreen question of the place of lectures in a university syllabus. Soon I was led to consider a thought-provoking discovery. I had come to know as friends and to admire three professors, world-famed for their advancements or knowledge, but regarded by students as unsuccessful lecturers. One was very nervy and edgy before facing his large and rowdy class, another, as finicky as Henry James about verbal distinctions, would keep an amused audience waiting while, by audible trial and error, he sorted out the right word. To the third, lectures were, perhaps, like fleas on a dog; he brushed them off as irrelevant distractions to his trains of thought, and would occasionally interrupt his free associations with "Let's see, where was I, Miss X?" All received the Order of Merit.

It was sometimes said that at one university parents cheerfully paid good money to tutors who advised undergraduates which lectures or lecturers, to avoid. Why not? Pilots know the whereabouts of sandbanks. The analogy is scarcely perfect, since it has been said, and in print, that the tutor's opinions of a particular sandbank's dangers were based on hearsay, even gossip. One read and heard that specific if less ruthless advice was given to the university student. He was told to avoid most lectures (Boswell writes that Dr. Johnson said this, but Boswell often Boswellised his own reports) to buy relevant books and to study them in the privacy and quiet of his own room. The student was assumed to have enough money to buy the books and rent a room, but knowledge of the development rate of the student's main and subsidiary subjects ought to have guided the adviser's judgment. Sources mentioned by a good lecturer may be in a foreign language and in journals not easily accessible. Some of the information he imparts may be unpublished, since keen lecturers who attend congresses often consider it a pleasant duty to pass on good news.

The alleged and real uses of examinations were often discussed especially by teachers familiar with procedures in several universities in their own country and in others. One heard of academic tipsters famed for precognition of questions, "almost certain to come up," and of examiners, often young, who chortled "Not those I shall set!" But then they might be told the tale (not as mythical as some think) of the old hand who said wearily to a tyro, setting a question from each of a heap of books: "In this place, whatever questions you set, you get the same answers."

The functions of the external examiner were considered. Some university departments invite him to set, with knowledge of the syllabus, a fair number of questions, not enquiring whether their themes have been lectured upon in detail. The aim is to give a first class examinee a chance to show originality, or evidence of having read more than prescribed chapters in prescribed books.

Psychologists would find it interesting to discuss the qualities desirable in a university or college lecturer, as distinguished from a tutor, coach or school teacher. How far is he or she actuated by motives which make for good acting? An actor speaks lines written by someone else, perhaps centuries ago and accepted as impressive in their own right. If early success has caused him to be typecast for years, he may even speak lines which nauseate him. His tones, pronunciations and gestures may have been knocked into him by an overbearing producer. Each performance, good or bad, finishes with the drop of the final curtain.

In contrast, a lecturer speaks his own lines, seldom learnt by heart. A good lecturer often uses notes unobtrusively, and may be amused at the adulation often given to a speaker who scorns the safety-net of notes. The lecturer need not watch anxiously after each paragraph for audience-reaction. He may have realised that while 10 per cent. of his hearers can be bored, even resentful, another 10 per cent.

or more may be interested, desirous of hearing the subject put in perspective, even if they criticise this, and regard a section which they did not grasp entirely as meriting a visit to the library. He may learn that the average student's view of a successful lecture can vary from country to country, even from university to university. I am told by a lecturer with experience in an old and a new university in England, and in several in America, that 'over there' students may expect to hear succinct arguments 'for' and 'against' with an authoritative summing-up. Not for certainty, they despise the mugwump. Yet certain ex-students say that they enjoyed spotting the hurdle which the lecturer was approaching, and seeing how, and if, he cleared it.

For some readers, the committee's most disturbing jolt is :

... "University buildings and plant and the time of university teachers will have to be better used. This makes the need for operational research and experiment in university teaching still more pressing. Such experiments should be conducted on an inter-university basis."

BOOK REVIEWS

CARPENTER, FINLAY, and HADDAN, EUGENE E. (1964). *Systematic Application of Psychology to Education*. New York : Macmillan, pp. vii+269, 22s.

GREEN, DONALD ROSS (1964). *Educational Psychology*. Foundations of Modern Psychology series. New Jersey : Prentice-Hall, pp. v+120, no price.

These two paperbacks have some features in common. Both are designed to meet student needs and finance, both are concerned about the failure of education courses to assist the student to acquire teaching skills, and both depart from the traditional textbook treatment of educational psychology. Their orientation are different, however.

Professor Green's book classifies and examines the factors relevant to school learning : (a) factors in school situations that influence learning processes ; (b) some of the major features of these learning processes ; (c) instructional procedures ; and (d) the effects of schooling. It is one of the very useful 'Foundations' series, but is less successful than some of the others because it is much less a specialist text than condensed general text book. It attempts to cover too much ground for its limited format and the author has obviously had a problem in choosing illustrative studies. Nevertheless, it is clearly written, avoids over-technical language and should provide useful supplementary reading for students in training.

The same cannot always be said of the book by Carpenter and Haddan, the usefulness of which, for British students, is severely limited by the assumption that students will already have some acquaintance with psychology and in particular, Skinnerian terminology. This is a pity, because it contains a rigorous analysis of 'Education in Subject Matter' in a way which is new and stimulating. The authors attempt to describe the instrumental behaviour of both teacher and learner in relation to the activities of information getting, application analysis and creativity. Essentially, it is an attempt to describe classroom behaviour in Skinnerian terms with a great deal of emphasis in defining objectives and applying criteria. Part three of the book is devoted to the analysis of some teaching methods, namely, the lecture, group discussion, the use of motion pictures, field trips and automated devices. Students are likely to find this the most interesting part of the book and it is certainly material that they will not find elsewhere. The final chapter contains some controversial recommendations about the relationship between teaching and research in schools of education in America which are also relevant here.

P. K. POPPLETON.

CARTHY, J. D., and EBLING, F. J. (1965). *The Natural History of Aggression*. London : Academic Press, Inc., pp. 159, 30s.

This book consists of the proceedings of a symposium on aggression which was held at the British Museum in London in October, 1963. Aggression is considered from the separate angles of such disciplines as zoology, physiology, psychology, psychiatry and sociology. The editors did well in attempting in the introduction to find common ground among the different contributions. It would have been to the book's advantage if a subject index had been provided and references systematised.

Certain basic questions loom large in several of the individual contributions and in the discussions. One is whether aggression is a universal 'interspecific' phenomenon. Wallis points out that both intra-species and interspecific aggression are needed to maintain colony cohesion in social insects. On the other hand, Fisher considers—and Lorenz concurs—that in birds and mammals, apart from predation and defence, aggression is essentially an intraspecific phenomenon. Furthermore, within any species, aggression to the point of killing is very rare : 'thrashing' and 'ritual' fighting are best attuned to the biological ends of producing fit individuals, spacing them out, and achieving functional social hierarchies.

A related question with regard to human beings is whether aggression is a necessary or merely a readily acquired mode of behaviour. Some contributors, e.g., Burton, regard aggression as mainly acquired, while others, e.g., Harrison Matthews, seem to incline to the opposite view. What impresses the reader is that the contributors examine and weigh evidence carefully in considering which forms of aggressive behaviour, and in what measure, are genotypic or otherwise.

The word aggression is, of course, not unambiguous; and Veness rightly attempts to re-examine and clarify concepts such as aggression and hostility. In dealing with aggression, the biological and the social scientists barely talk the same language. The contributions of the latter are not altogether overshadowed by those of the former. Andreski's discussion of the origins of war is most interesting. No less so, but very peripheral to the symposium, is a paper by Laver on costume as a means to social aggression. It would be possible to single out several other contributions as especially interesting and valuable. Suffice it to say, however, that *The Natural History of Aggression* is a useful book, and one of great interest to all concerned with comparative psychology, in the broadest sense of this term.

W. SLUCKIN.

DEMBER, WILLIAM (1964). *Visual Perception—The Nineteenth Century*. New York: John Wiley, pp. xii+222, 38s., 23s.

In the paper-back series, *Perspectives in Psychology*, edited by Kessen and Mandler, Professor Dember presents a series of excerpts from nineteenth century writings on visual perception. It is to be followed by a second volume of contemporary material on the same subject. This book is designed to interest students of the history of psychology as well as of perception. It is too advanced for the beginner in psychology; but those who are already familiar with modern work, on perception will be interested in studying the early stages of theory and observation on this subject, and in reading about the origins of various concepts of which they have learnt only through the writings of later authors. Students may also be stimulated to undertake further reading in the original sources, though more information might have been included to indicate such reading.

Since the excerpts are confined to the nineteenth century, Professor Dember cannot begin, as he might, with Berkeley, although he mentions him in the introductory material. The first section, containing excerpts from early writings on psychology, does not appear very relevant to the later work on visual perception. Much the most interesting selections are those on visual illusions and on binocular space perception, especially Wheatstone's studies of the stereoscope. Later sections are rather disappointing. Throughout, the reader may be struck by the discursive manner of these writings and the 'field naturalist' approach to perceptual phenomena. It is regrettable that no experimental studies are included, such as those of Wundt and Cattell, to demonstrate that experimental investigation of perception did, in fact begin during the nineteenth century.

However, Professor Dember must be commended for the ingenuity of his introductory and linking discussions, though he should perhaps have made it clearer which are verbatim transcriptions from the originals, and where there are omissions and abbreviations. Though the book may appeal in the main to specialized interests, it is nevertheless, of real value.

M. D. VERNON.

DUVAL ROCH (1964). *Adolescence D'Aujourd'hui*. Quebec: Les Presses De L'Universite Laval, pp. 191, \$3.00.

The phenomenon of adolescence in any society always involves the question of how far what happens is the outcome of maturation and development and how far it is the consequence of cultural influences. From this aspect, *Adolescence D'Aujourd'hui* is of particular interest. It brings us news of adolescents acquired from the French Catholic region of Canada and presented by a writer with this cultural background. Familiar sources of stress emerge: desire for independence from parents and in the family; desire for social independence; desire for economic independence, and the desire for heterosexual love. One might define adolescence

psychologically, it would seem, by the presence or absence of these areas of stress and conflict.

But this book does more than cover again the well-known ground. Roch Duval shows originality by adopting an existentialist attitude to his theme, drawing confirmation not by reference to statistics, but by quotations from the novel, *The New Aristocrats*, by Michel de Saint-Pierre, himself a practised observer of adolescent behaviour. Such a method can easily degenerate into mere anecdotalism but, as used by the author, the method permits the un inhibited statement of acute descriptions that ring true because they carry the authenticity of objective, sympathetic observation. To translate, and paraphrase, some of these insights: "What adolescents most need is CALM, for they have perpetual tensions," "Conflicts are not necessarily bad—indeed, often essential—but indissoluble conflicts are bad; the ego withdraws." "The violence of an adolescent's demands simply indicates his need to be taken seriously." "Adolescents are both highly-sexed and modest." "Don't imagine that between adolescents simple friendship is enough to complete the education of the affections." As might be expected from a religious moralist, the author lays much stress on control, and takes the reader into such unexpected areas as the correlates of religious vocation, but he sees control as the consequence of individual maturation, not as another name for inhibition. Throughout, his outlook is developmental: "Educating adolescents to have a healthy ego will produce better Christians than old-style religious teaching"—a conviction pertinent to the problem of moral education everywhere.

The second part of the book presents a critical appraisal of secondary education in the light of the realities of adolescent development. The author wishes to see more attention paid to the emerging personalities of secondary school children, more group discussion; less spoon-feeding, less homework. Here, too, experience elsewhere returns an echo.

JAMES HEMMING.

GOLDMAN, RONALD (1964). *Religious Thinking from Childhood to Adolescence*. London: Routledge and Kegan Paul, pp. xii+276, 32s.

There are many problems in the highly controversial field of religious education. First, and most fundamental, is whether any attempt at all should be made in publicly supported schools to teach children religious knowledge, when there are many who would strongly call into question the very credentials of such 'knowledge.' Second, granted that some such attempt should be made, what particular brand of religious ideas and concepts, of the many available, should be taught to the children? And third, given agreement on the form and content of 'true' religious knowledge, how might this most effectively be expounded and explained to children of different ages, abilities and backgrounds?

Only the third of these questions is a matter on which psychologists, as psychologists, may be expected to pronounce. It is with such a problem of applied psychology that Dr. Goldman is principally concerned in this book. Certainly, answers to the earlier, and more basic, questions are assumed: religious knowledge should be taught in the schools, and such teaching should be consistent with modern theology and biblical scholarship. But these matters are not argued at any length. Rather, the author's main purpose is to give "a descriptive account of how school pupils think about religion," with the intention of thereby improving our methods of religious education. He maintains that religious thinking is no different, in principle, from other sorts of thinking, and thus that psychological accounts of thinking in general may validly be used to explain religious thinking. Dr. Goldman's preferred conceptual framework for understanding children's thinking is that of Piaget. The results of an extremely detailed and devoted enquiry into the religious ideas of a sample of 200 school children of varying ages, abilities and backgrounds are presented at length, and shown, with some plausibility, to be interpretable in terms of Piaget's developmental structure of thinking. More especially, it clearly and repeatedly emerges that much present-day teaching of religious knowledge is wildly unsuitable for those of the age-levels and abilities to whom it is directed;

and that early acquired misconceptions, for example, interpreting many biblical stories as literal historical truths, may play havoc with later attempts to understand religious ideas in a more sophisticated and scholarly way.

Dr. Goldman's book is, therefore, strongly to be recommended to all who are interested and involved in problems of religious education. True, there is little here for those who reject the whole notion of religious education in the schools, or for those, such as fundamentalists, whose interpretation of religious knowledge is vastly different from that of the author and of modern theologians. Others will regret the exclusion from the enquiry of "children of parents who were Roman Catholics, Jews, negroes or of foreign extraction." Nevertheless, here is a most valuable and illuminating empirical study of the actual struggles of children of varying ages and abilities to understand complex and often symbolic religious ideas in a world increasingly dominated by logical-scientific thinking.

G. W. PILKINGTON.

KERLINGER, FRED. N. (1964). *Foundations of Behavioral Research*. New York : Holt, Rinehart and Winston, pp. xix+739, no price given.

Intended for "graduate students who have elementary backgrounds in psychology, statistics, and measurement" (p. ix), this is a big book both in conception and execution. Its eight major sections are: (1) the language and approach of Science; (2) sets, relations, and variance; (3) probability and statistical inference; (4) designs of research; (5) types of research; (6) measurement; (7) methods of observation and data collection; and (8) analysis and interpretation.

Parts I, II and III develop concepts for use in later discussions. The core vocabulary is economical and is reinforced by use throughout the remainder of the book. A good symbolism is developed in Part IV for comparing research designs which ably expresses the difficult point about whether or not the independent variable is manipulable, and if manipulable whether it is or is not manipulated. This must be one of the clearest accounts of elementary design principles that is available today. Part V discusses research programs largely according to whether the independent variable is currently manipulable as in laboratory experiments, or is derived *ex post facto*. The properties of different levels of measurement and the logic of test measurement are discussed with clarity and economy of concept in Part VI. Part VII reviews the interview, objective tests, observational methods, projective methods, content analysis, sociometry and includes very full chapters on the Semantic Differential and Q Methodology. The first two chapters of Part VIII discuss descriptive methods, tabulation and the like; and a final chapter explains factor analysis. Three appendices cover the writing of research reports, historical and methodological research and the use of computers in the behavioural sciences.

Many chapters close with well-chosen readings for further study. Throughout the book a rich background of references is evident in numerous footnotes. Unfortunately, the brief author index is far from complete and the reason for excluding some authors is not evident. The research examples are always apt and their descriptions neatly trimmed so that the relationship of the substantive hypothesis to the research design and its outcome is underlined. Hypothetical numerical examples are expressed by single digit numbers as a matter of policy, and real data from the research literature are presented simply. There is an admitted bias towards educational research, but this avoids some of the difficulties that might have resulted from the use of specialist psychological terms.

Following the careful progression of concepts in the first two parts, Part III, the statistical core and largest section of the book, is disappointing in some respects. First, chi-square is introduced precipitously, acting merely as a vehicle for discussing statistical significance (p. 150). For this reason, the student is asked to take a number of things on trust quite unnecessarily. The null hypothesis model is in any case stated in terms of coin-tossing and could have been followed through as such. Secondly, the concept of "degrees of freedom" is hurriedly and inadequately introduced, although one or two less fundamental considerations are given space. Treatment of the normal distribution is also hurried, and standard scores are explained in

a footnote (p. 163). Thirdly, having carefully laid the foundations for a discussion of components of variance, it is surprising to find the chapter on the t test preceding those on analysis of variance. The formula given for the standard error of the difference between two means (p. 177) is not, moreover, in the form that makes clear the derivation of t from the F ratio. Finally, there is some confusion of symbols for variance, mean squares, and sums of squares across pages 189 to 197. One symbol stands both for the 'variance of the means' and the 'between groups variance' with the consequence that numerical results differ by a factor of n , the size of the group. Expressions for the 'sums of squares between groups' also differ by the same factor on pages 190 and 197. All this is probably due to the insertion of the t test discussion.

It would, however, be churlish to let these criticisms detract from the overall competence of the book. With its stable-companion, Hays' *Statistics for Psychologists*, it clearly belongs to a welcome new generation of text books in this field.

PHILIP LEVY.

KLEIN, JOSEPHINE (1965). *Samples from English Cultures*. London: Routledge and Kegan Paul, 2 Vols., pp. xii+334, xi+334, Vol. I 50s., Vol. II 35s.

These two volumes are about social influences on personality development. Studies of English life, which have appeared in the last decade, are used to illustrate (mainly American) propositions about connections between personality and social experience. Descriptions of family and other social relationships are tied in with American interpretations of the achievement motive, duodenal ulcers and the bureaucratic personality.

The books quoted and summarised are numerous and diverse. This is one of the difficulties. Hoggart and Gorer carry as much weight as Stacey and Mosey. Spinley, Kerr and the Newsons shared Dr. Klein's interest in personality and 'culture'; most of her other sources, when they are not higher journalism, are concerned primarily with social structure.

Personality can be examined to explain, in part, the origins and character of a culture; and culture can be examined to explain, in part, the shaping and perpetuation of personality. The 'cultures' of Dr. Klein's title apparently call for no explanation, whether in psychological or socio-logical terms. This is, perhaps, consistent with her claim merely to describe what is there, to offer an 'essay in social cartography.' But the reader is entitled to know what these things are that are being located in social and psychological space. He is not told. And he is given little help if he feels inclined to guess.

We are given a definition (from the Shorter Oxford) of a 'sample.' We might have managed this for ourselves. But we get no help with cultures (or sub-cultures). And this we all urgently need. This ubiquitous term carries a diversity of meanings, its imprecision plagues attempts at analysis in the social sciences. In particular, we need to know whether we are talking about a shared normative system or simply statistically frequent behaviour. There is evidence throughout these two volumes of changes in behaviour when the situation changes, for example through migration to new housing estates. Social-class differences in levels of educational aspiration which are cited yet again, may in part, be situationally rather than normatively induced. If our understanding of these 'cultural' differences is to be advanced, we need to know whether similarities in behaviour drive from common values or simply from common opportunities.

There will always be a relationship between culture and personality; but the relationship is not always the same. If we are considering a 'contraculture,' which arises from a situation of conflict, there will be one kind of linkage; if we are describing a subculture which arises, for instance, from geographical isolation, the linkage will be different. Delinquent gangs and adolescent peer groups are subcultures which have one kind of linkage; isolated rural or mining communities have another. The Paddington studies which are recapitulated here are suggestive of a contraculture, at odds with wider society, beleaguered, living by inverse values; the Yorkshire mining community is perhaps more properly a sub-culture, to be explained

in sociological terms. An ethnic enclave is often such a subculture, satisfactorily explained in social terms of migration, isolation, occupation. Such personality tendencies as anxiety, inadequacy, and resentment need play no part in its origin or continued existence.

Dr. Klein uses a variety of sociological concepts to explain subcultural behaviour—from anomie and relative deprivation to role. This is sometimes illuminating; but without proper clarification of concepts can lead to confusion. The subculture of a group distinguishes and separates it from the wider society; roles do the opposite. Miners and adolescents may have a subculture which separates them from the wider world; but the 'strictly defined' roles of miners and adolescents, to which Dr. Klein refers, relate them to society, so that people outside their subcultural groups know what to expect of them. Of course, there can be specifically subcultural roles; but we do not further sociological or psychological analysis by using role and culture as if they meant the same. "The theoretical background is uncomplicated," Dr. Klein believes. It is far more complicated than she knows.

F. MUSGROVE.

MARKLE, S. M. (1964). *Good Frames and Bad: A Grammar of Frame Writing*. New York: J. Wiley, pp. xiv+278, 35s.

This book considers one of the several stages that are undergone in the production of a programme, namely that of frame-writing. It is important to observe that it is not intended as a full treatment of programmed instruction. This *Grammar* is written for students who are interested in programming techniques and who are high in verbal and scholastic ability: it is not designed to produce professional programmers. The book contains four chapters, three on styles of programming (linear, multiple-choice branching, and—very briefly—mathematics) and the fourth on editing. The student experiences a wide variety of programming styles, and a host of useful examples and procedures: he is constantly engaged in a critical analysis of programming techniques and theory.

The layout of this programmed text is like no other known to the reviewer: it is essentially an intelligent format for adult readers. Miss Markle happily scotches the idea that all the frames in a programme should be the same physical size, and reiterates what Skinner argued all along—each step should be such a size that the student can reasonably expect to master it. With such a population, this means that the frames vary in size from the minute to the extremely large. Nor are these frames of the 'Swiss-cheese' variety, full of holes or blanks to complete—conversational questions are asked, opinions are given and discussed, and only occasionally is a written response required. In addition, a heavy and a pleasing use is made of the response element of the text for confirmation of the students' response, and for further discussion and elaboration. One limitation of the programme is, however, that it is too much concerned with individual frames. In programmed instruction stimuli are presented in sequences. Unfortunately, the author has seen fit to omit any reference to her earlier excellent paper on the analysis of effective sequences in programmed instruction (*vide: Programmed Instruction*, III, 1, pp. 4 and 5, 1963).

Programmed Instruction is ten years old. During this period Susan Meyer Markle has contributed a series of papers which have always stimulated readers. It is now, however, an appropriate time to ask what progress has been made. It would seem, despite the qualities of this book outlined above, that there has been, in fact, little progress in frame writing. Miss Markle has excellently put into practice in quite the same way as her previous work. Perhaps the reason for this is that one is expecting too much: this book is not about programmed instruction, it is about frame writing. Frame writing is not the core concept or difficulty in programmed instruction. Subject matter analysis and content sequencing are more important aspects where definite progress has been made. *Good Frames and Bad* makes an important contribution, but to a less important aspect of the field.

JAMES HARTLEY.

NIBLETT, W. R. (Edit.) (1965). *How and Why Do We Learn?* London: Faber and Faber, pp. 196, 21s.

Many a good public lecture deserves a wider audience, though some lectures survive the process of printing more successfully than others. In recent years, several series of lectures on educational topics given in the University of London have been published in book form. In each series, however uneven the quality, we can usually be sure of finding at least one or two lectures which themselves justify the publication.

This text is a series of lectures on educational psychology, delivered in 1963-64. Aspects of educational psychology have been skilfully divided among the contributors. The result is a book which is interesting as a unit, while individual chapters will provide a student with a useful review of a selected theme.

In the first lecture, "Learning *versus* Teaching," Professor Stephen Wiseman pleads that in education our preconceptions should be tested against evidence, and proceeds to review research on a variety of relevant topics—maturation, environmental influences on development, teachers and teaching methods—to illustrate his point. Next, Dr. Doris Lee discusses the nature of perception, intuition and insight, and the relevance of research on these to education. Dr. W. D. Wall, in *Learning to Think*, shows the relevance of learning theory to educational practice. Professor Harry Kay discusses some problems in programmed learning. Professor Ben Morris explains the dynamics of small group situations, considers leadership, and draws the implications for education. Dr. W. Taylor continues the account of social learning, extending the scope to include the school and the neighbourhood.

There is also a Part Two, with three further lectures: Richard Hoggart on the mass media; Stephen Potter's witty *Learning to Enjoy*; and an autobiographical account by Sir Hugh Foot of his diplomatic activities, entitled *Learning to Live in a Heterogenous World*. In the Introduction the editor defensively seeks to justify the inclusion of these extra contributions, perhaps anticipating the reader's (and the reviewer's) misgivings.

The variety of treatment is such that most readers will find at least one of the lectures stimulating, and it is unlikely that anyone will enjoy all of them. Of course the advantage of a book over a lecture series is that it is easier to move on when interest flags.

A similar series of lectures published in 1955 (including Knight, Fleming, Morris, Wall, Vernon and others) was entitled "The Bearings of Recent Advances in Psychology on Educational Problems." The 1965 title is more striking. But it is interesting to note that the optimistic note of 'recent advances' in 1955 has been replaced in 1965 by a question mark.

JOHN NISBET.

PHILLIPS, MARGARET (1965). *Small Social Groups in England*. London: Methuen, pp. 318, 25s.

This is an odd book out in the voluminous literature on groups. It covers an extremely wide range, from small and temporary informal groups to segments of large institutions such as the Services, hospitals and schools, and the time span extends over a period of some twenty years. This cannot be achieved by conventional methods, and the author's chief source of material was a questionnaire handed out to group members asking them to record their observations and experiences. Out of the mass of data thus assembled, Miss Phillips has distilled a series of variegated vignettes of group life; only those who have had to deal with this kind of material will be able to appreciate the immense amount of labour entailed in such a task. The process of selection was guided by a theoretical framework which is a somewhat eclectic mixture of Freud, Talcott Parsons and others. At the end of each chapter, featuring sketches of several groups, there is a discussion designed to bring out theoretical issues connected with the material.

In contrast to many dreary tomes on groups, this book makes easy reading ; the observers often write with gusto, and evident feeling. At the same time it also becomes apparent that many of the observers were far from detached, judging their fellows from a particular standpoint : here is an example : " There was, however, a left-wing tendency among the younger members, a more balanced outlook among the older." Apart from religious, political, social class and other preconceptions, a person's perception of a group may also be strongly coloured by his special position within it. While the author is, of course, aware of these factors, she would seem to make insufficient allowance for them, often tending to write as though the account were a purely objective one.

Another shortcoming also arises inevitable from the method used and from the way in which the work took shape. The original collection of data was innocent of any particular theoretical orientation—indeed, at the time very little worthy of that label existed. Hence, theories came in at a relatively late stage, the data being used to illustrate them, especially a rather simplified ' family ' model of a psychoanalytic type ; moreover, the author appears to be unfamiliar with a good deal of recent work on group behaviour and leadership, such as that of Newcomb, which would be highly relevant to her thesis and might have led her to change some of the emphasis. Thus, for instance, she cites with approval the statement by Homans that the leader must be the most conformist in the group, interpreting several of the group settings in terms of this alleged principle ; but Hollander has shown convincingly that the statement in this rigid form is untenable.

In view of the author's modest claims these criticisms might seem churlish, but it was felt advisable to sound a cautionary note lest anyone be misled into trying to use this book as a text on group behaviour. Miss Phillips merely suggests that her work might be a source of hypotheses, particularly concerning the development of groups in time. Further, as Bryan Wilson puts it in " A sociologist's footnote " (a very informative chapter added to the book) : " One obtains a particularly vivid picture of the self-interpretation of groups and of their self-awareness." This is a dimension lacking in most contemporary work, and here the author has made a distinctive contribution. Educationists will probably find the chapter on " Teaching staffs " rewarding ; teachers concerned with group behaviour can use the book as a source of case histories on which students can sharpen their analytical tools.

GUSTAV JAHODA.

SLUCKIN, W. (1964). *Imprinting and Early Learning*. London : Methuen, pp. x+147, 25s.

The phenomenon of imprinting has been known about for a long time. Spalding first described it in 1873, and Lorenz gave it scientific respectability in his famous paper, " The companion in the bird's world," in 1935. Yet it is only in the last few years that imprinting has aroused the curiosity of psychologists and become a ' fashionable ' area of research. National and language barriers, the intervening war years, and the inevitable lag required to digest any new finding may have had something to do with this, but more likely the sudden interest in this topic stemmed from the growing dissatisfaction with orthodox reinforcement theory and the realisation that in the gosling's learning to follow the mother goose in the absence of food rewards, and, apparently all the more enthusiastically, despite punishment, a challenge to theoretical ingenuity and adaptability was held out. As a result, the research literature on this topic has suddenly swelled and, because of its scatter in zoological as well as psychological journals, made a summary of findings and a gathering together of the theoretical threads highly desirable.

In this book Sluckin has undertaken this timely task. He has searched the literature assiduously and drawn up a most readable account of the present state of knowledge and thinking about imprinting. Most of the work he describes has been done on birds, for here the following response assumes a clear-cut form and more sparse, from various mammalian species indicates that imprinting-like behaviour is also to be found higher up the evolutionary scale in such animals as

goats, sheep, dogs, deer, buffaloes, and zebras, and it is, therefore, inevitable that the implications for human development come up for consideration. On this point, however tempting it may be to indulge in speculation, Sluckin is careful to be cautious, for research on early human learning is still strangely retarded when compared with work on ducklings, goslings and chicks. Only the notion of maturational readiness for certain types of learning experience, which underlies the imprinting phenomenon and which has been expressed in the concept of critical (or, more cautiously, sensitive) periods, is one that has wide applicability, and a book such as the present one may well be read with profit by those who are inclined to discuss the learning process in isolation and refuse to take into account its developmental context.

While much of the book is devoted to an examination of the variables affecting the manifestation of imprinting, Sluckin by no means neglects the wider theoretical issues. Do the findings discussed under this heading show that there is more than one kind of learning, does early learning differ quantitatively from later learning and, in particular, must the reinforcement principle be supplemented or even replaced by other types of explanations. Sluckin raises the issue whether the study of imprinting has not brought up the need for a new theoretical formulation for learning and motivation, and though he does not undertake such a mammoth task himself, here his discussion of imprinting as a form of "exposure learning" and the linkage he makes with perceptual learning may well prove a step in just this direction. It is somewhat frustrating that the discussion of this highly interesting motion is so brief and that the theoretical analysis and integration of empirical findings should not have been carried further, for this is really the climax of the book and, in many respects, potentially its most important contribution. One must hope that the author will develop this notion elsewhere.

It should be pointed out that the dust cover is rather misleading when it states that the author, in his conclusions, discusses the implications which imprinting has for educational and abnormal psychology. These receive but a bare mention, for Sluckin quite rightly eschews the temptation to indulge in the unwarranted extrapolations and generalisations to which imprinting has been particularly subject. But to anyone who wants to follow the development of a new and fascinating area of behavioural research and to witness the extension of one particular frontier of knowledge, this book can be highly recommended.

H. R. SCHAFFER.

VERNON, P. E. (1964). *Personality Assessment: A Critical Survey*. London Methuen, pp. ix+333, 42s.

Some giants do not bother much about the work of people of lesser—or even similar—stature. Philip Vernon has never been in this group. Writing twenty-five years ago about his *The Measurement of Abilities*, the reviewer remarked: "Not the least of the services he performs is the preparation of statements of the problems, methods and findings of other psychologists." Professor Vernon is still at it; and those whose work he scrutinises should feel gratitude for the meticulous attention their efforts receive from him. True, his comment is sometimes quietly devastating, but never without apparent justice.

Occasionally, perhaps, the comment is less effective than it could be. For although Dr. Vernon's erudition flows on and on, it has its jerky patches. Research findings of people you have never heard of are fascinatingly teased out and hitched together; and just as you are expecting some of the 'buts' to be explained, or some new insight to be presented, oft you go to something else. Vernon makes few concessions to people who like bridges: he thinks stepping stones should be enough. Perhaps part of the explanation is to be found in an old-fashioned remark made during the war by one of his former colleagues: "Philip doesn't realise that some of our g's wouldn't come above his ankles."

Sometimes he provides a surprise. For instance, he still ignores what some of us think was a very nasty flaw in the Kelly and Fiske study of predictors and criteria of performance in clinical psychology—the artificiality of the whole

situation. How can you claim to be seriously engaged on the validation of selection tools, when you know that your subjects have known, from the beginning of the 'selection' procedure, that they are already home and dry? The whole set-up was breath-taking in its naivete. Yet Professor Vernon's criticism of it shows no real advance on the one he put forward in his *Personality Tests and Assessments*, which came out not long after the Kelly-Fiske report and was, therefore, in some degree excusable.

Again, he displays an odd persistence, right from the first page of his foreword, through chapter one, to the end of the book, in attributing 'Freudian or depth' ideas to clinical psychologists *as* clinical psychologists. He does not quite say that this is part of his definition of clinical psychology; but he gets very near to it, through his repetitious conjunction of these adjectives. Thus seems a little perverse.

But such blemishes as these cannot be regarded as desperately important, for the book is without doubt a major psychological publication. It is a pity that he has hidden away in various places (though it is not so hidden in Chapter 11) what he has to say about Cattell and Eysenck, who are both good for several lines of page numbers in the author index. (Their runners-up are Guilford whose views were shared 'at one time' by the author followed by Cronbach and Gordon Allport). The discussion on Cattell and Eysenck goes a good deal deeper than the splendid chapter on factor theories provided eight years ago by Hall and Lindzey, in their *Theories of Personality*; but it does, perhaps, need the Hall-Lindzey exposition (or something like it) as briefing material. And perhaps it should have, as supplements, two papers ('Eysenck on Cattell' and 'Cattell on Eysenck') which appeared in *Occupational Psychology*, in 1961 and 1964, respectively.

ALEC RODGER.

PUBLICATIONS RECEIVED

The mention of a book in this list neither implies, nor precludes, a later review.

- OTTO L. SHAW (1965). *Maladjusted Boys*. London : George Allen and Unwin, Ltd., pp. 168, 28s. net.
- K. M. EVANS (1965). *Attitudes and Interests in Education*. London : Routledge and Kegan Paul, pp. viii + 168, 16s. net.
- ROBERT I. WATSON (1965). *Psychology of the Child*. New York : John Wiley, pp. ix + 635, 57s.
- HENRI TAJFEL and JOHN L. DAWSON (1965). *Disappointed Guests*. Lond: Oxford University Press, pp. vii + 158, boards 25s. net, limp 15s. net.
- MICHAEL YOUNG (1965). *Innovation and Research in Education*. London : Routledge and Kegan Paul, pp. viii + 184, 25s.
- HAROLD W. BERNARD (1965). *Psychology of Learning and Teaching*. New York : McGraw-Hill Book Company, pp. xix + 539, 64s.
- RAYMOND HICKEL (1965). *Modern Language Teaching by Television*. Strasbourg : Council for Cultural Co-operation of the Council of Europe, pp. 185, 10s. sterling.
- ROGER REGER (1965). *School Psychology*. Illinois : Charles C. Thomas, pp. xv + 213, \$7.75.
- PETER B. NEUBAUER (1965). *Concepts of Development in Early Childhood Education*. Illinois : Charles C. Thomas, pp. ix + 149, \$6.50.
- D. S. LEHRMAN, ROBERT A. HINDE, and EVELYN SHAW (1965). *Advances in the Study of Behaviour*. Vol. 1. New York : Academic Press, Inc., pp. x + 320, 76s.
- (1965). *L'Annee Psychologique*. Paris : Presses Universitaires De France, pp. 718, no price.
- (1965). *Revue Roumaine Des Sciences Sociales*. Bucarest : Editions De L'Academie De La Republique Populaire Roumaine, pp. 136, no price.
- THOMAS E. CLAYTON (1965). *Teaching and Learning*. London : Prentice-Hall International, pp. xi + 177, 18s.
- J. E. MEADE and A. S. PARKES (1965). *Biological Aspects of Social Problems*. Edinburgh : Oliver and Boyd, Ltd., pp. x + 226, 50s.
- W. G. WALKER (1965). *The Principal at Work*. Australia : University of Queensland Press, pp. xi + 219, 37s. 6d.
- PAUL GYORGY and ANNE BURGESS (1965). *Protecting the Pre-School Child*. London : Tavistock Publications, pp. xviii + 108, 18s. net.
- MAURICE DE MONTMOLLIN (1965). *L'enseignement Programme*. Paris : Presses Universitaires De France, pp. 128 + 8, no price.
- F. H. GEORGE (1965). *Factors*. Bristol : Teaching Programmes, Ltd., pp. xiii + 23, 6s. 6d.
- J. H. HORROCKS (1965). *Assessment of Behavior*. U.S.A. : Charles E. Merrill Books, Inc., pp. xv + 736, 68s.
- K. C. GARRISON (1965). *Psychology of Adolescence*. U.S.A. : Prentice-Hall, pp. xii + 487, 64s.
- WM. CLYDE DEVANE (1965). *Higher Education in Twentieth Century America*. U.S.A. : Harvard University Press, pp. vi + 211, 36s.
- ELLIS BATTEN PAGE (1965). *Readings for Educational Psychology*. New York : Harcourt, Brace and World, Inc., pp. xii + 404, \$3.95.
- J. G. WALLACE (1965). *Concept Growth and the Education of the Child*. Bucks : National Foundation for Educational Research, pp. 268, 42s.
- ARTHUR G. HOLT (1965). *Handwriting in Psychological Interpretations*. U.S.A. : Charles C. Thomas, pp. xiii + 26, \$10.50.
- OLENA SENNTON (1965). *Exposure to Films and School Adjustment*. Stockholm : Almqvist and Wiksell, pp. 287, no price.
- D. R. MATHER, N. FRANCE and G. T. SARE (1965). *The C.S.E.—A Handbook for Moderators*. London : Collins, pp. 188, 15s.

- M. L. KELLMER PRINGLE (1965). *Deprivation and Education*. London : Longmans, pp. vi+311., 42s. 6d.
- R. L. CURLING and A. BLAKELEY (1965). *Echoes of Experience*. London : Longmans pp. x+86, 5s. 6d.
- WILLIAM J. MEYER (1965). *Developmental Psychology*. London : Prentice-Hall, pp. 116, 36s.
- H. J. EYSENCK (1965). *Smoking, Health and Personality*. London : Weidenfeld and Nicolson, pp. 166, 25s.
- LEONARD P. ULLMANN, LEONARD KRASNER (1965). *Case Studies in Behavior Modification*. New York : Holt, Rinehart and Winston, pp. xii+401, no price.
- GEORGE F. KNELLER (1965). *Introduction to the Philosophy of Education*. New York John Wiley, pp. viii+137, 15s.
- CHESTER A. LAWSON and MARY A. BURMESTER (1965). *Programmed Genetics : Vol. II—Chromosome Behaviour*. Boston : D. C. Heath and Co., pp. vi+72 +A10, 22s. 6d.
- JAMES L. HENDERSON (1965). *Look Out*. Tent : New Education Fellowship, pp. 43, 5s.
- Z. P. DIENES and M. A. JEEVES (1965). *Thinking in Structures*. London : Hutchinson Educational, Ltd., pp. 126, 21s.
- PAUL SCHILDER (1965). *The Image and Appearance of the Human Body*. New York : John Wiley, pp. 353, 13s.
- HERBERT A. ROSENFELD (1965). *Psychotic States*. London : The Hogarth Press, Ltd., pp. 263, 42s.
- ANNA FREUD, RUTH S. EISSLER, HEINZ HARTMANN, MARIANNE KRIS and OTHERS (1965). *The Psychoanalytic Study of the Child*. London : The Hogarth Press, Ltd., pp. 493, 75s.
- R. C. JOHNSON and G. MEDINNUS (1965). *Child Psychology : Behavior and Development*. New York : John Wiley, pp. x+556, 53s.
- R. F. PECK and ROBERT J. HAVIGHURST (1965). *The Psychology of Character Development*. New York : John Wiley, pp. xix+267, 12s.
- C. L. HULL (1965). *A Behavior System*. New York : John Wiley, pp. ix+372, 17s.
- MAURICE MERLEAU-PONTY (1965). *The Structure of Behaviour*. London : Methuen and Co., pp. xxviii+256, 35s.
- JOHN STROUD (1965). *An Introduction to the Child Care Service*. London : Longmans, pp. viii+215, no price.
- COMMISSION OF INQUIRY ON MENTAL HANDICAP (1965). *Tuarascail Report*. Dublin : Stationery Office, pp. xxviii+195, 6s. 6d.
- ANDREA JADOULLE (1965). *La Psychologie Scolaire*. Paris : Presses Universitaires De France, pp. 163, F.8.
- HOWARD H. KENDLER (1965). *Basic Psychology*. London : Methuen, pp. xii+750, 60s. net.
- HENRI DIEUZEIDE (1965). *Les Techniques Audio-Visuelles*. Paris : Presses Universitaires De France, pp. 158, no price.
- D. E. BERLYNE (1965). *Structure and Direction in Thinking*. New York : John Wiley, pp. vii+378, 68s.
- D. H. STOTT (1965). *Roads to Literacy*. Glasgow : W. and R. Holmes, pp. 143, 17s. 6d.
- H. J. EYSENCK and S. RACHMAN (1965). *The Causes and Cures of Neurosis*. London : Routledge and Kegan Paul, Ltd., pp. xii+318, 28s.
- J. H. KAHN (1965). *Human Growth and the Development of Personality*. Oxford : Pergamon Press, pp. xvi+220, 25s.
- S. LEFF and V. LEFF (1965). *The Search for Sanity*. (Oxford : Pergamon Press, pp. v+146, 15s.
- F. B. DAVIS (1965). *Educational Measurements and their Interpretation*. U.S.A. : Wadsworth Publishing Co., Inc., pp. ix+422, 60s.
- BRIAN HOLMES (1965). *Problems in Education*. London : Routledge and Kegan Paul, Ltd., pp. x+326+13, 32s.
- DOROTHY CANFIELD FISHER (1965). *Montessori for Parents*. Massachusetts : Robert Bentley, Inc., pp. x+240, \$5.95.

- DOROTHY CANFIELD FISHER (1965). *The Montessori Manual*. Massachusetts : Robert Bentley, Inc. pp. 126, \$5.00.
- ERNEST FRIEDLANDER (1965). *Psychology in Scientific Thinking*. New York : Philosophical Library, Inc., pp. 143, \$4.85.
- REGINALD D. ARCHAMBAULT (1965). *Philosophical Analysis and Education*. London : Routledge and Kegan Paul, Ltd., pp. xii + 212, 28s.
- ROBERT K. YOUNG, and DONALD J. VELDMAN (1965). *Introductory Statistics for the Behavioral Sciences*. New York : Holt, Rinehart and Winston, pp. vii + 435, \$7.95.
- SIR GEOFFREY VICKERS (1965). *The Art of Judgment : A Study of Policy Making*. London : Chapman and Hall, Ltd., pp. 242, 25s.
- GEORGE A. C. SCHERER and MICHAEL WERTHEIMER (1965). *A Psycholinguistic Experiment in Foreign-Language Teaching*. London : McGraw-Hill, pp. xii + 256, 60s.
- MUSSEN, CONGER and KAGAN (1965). *Readings in Child Development*. New York : Harper and Row, Ltd., pp. x + 480, 37s.
- VICTOR H. NOLL (1965). *Introduction to Educational Measurement*. Boston : Houghton Mifflin Co., pp. xviii + 509, \$7.75.
- MILDRED B. KANTOR (1965). *Mobility and Mental Health*. U.S.A. : Charles C. Thomas, pp. xix + 247, \$10.50.
- JOHN M. STEPHENS (1965). *The Psychology of Classroom Learning*. New York : Holt, Rinehart and Winston, Inc., pp. xiv + 511, no price.
- ROBERT N. WILSON (1964). *The Arts in Society*. U.S.A. : Prentice-Hall, pp. xii + 372, 54s.
- EILEEN YOUNGHUSBAND (1965). *Social Work with Families*. London : George Allen and Unwin, Ltd., pp. 187, 25s.
- M. ESTHER HARDING (1965). *The Parental Image*. New York : G. P. Putan's Sons, pp. xviii + 238, \$6.50.
- F. G. LENNHOFF (1965). *Forty-four Children Thinking Aloud*. Salop : Shotton Hall Publication, pp. 52, 6s.
- ANDREW WILKINSON (1965). *Spoken English*. University of Birmingham, pp. 130, no price.

Articles :

Absence, the Effects of on Primary School Performance : J. W. B. DOUGLAS and J. M. ROSS	28
Achievement Motivation and Goal Setting Behaviour in the Classroom : BISHWA NATH MUKHERJEE	286
Attitudes and Attainment, Sex Differences in, in Junior Schools : M. WISENTHAL	79
Attitudes of West African Students and Teachers Towards the Modern Approach in Teaching, a Comparative Study of : OSCAR FERRON	294
Blind and Sighted Children, a Study of Personality Differences between : HAMED A. S. ZAHARAN	329
Co-education : The Verdict of Experience, II - A Qualitative Approach : R. R. DALE	195
Craftsmanship in Pottery Making at a Secondary School, Psychological Aspects of : G. H. JAMIESON	179
Creative Thinking, Effects of Restrictive Range and Test Unreliability on Correlation Between Measures of Intelligence on : KAORU YAMAMOTO	300
Environmental Handicaps and Intellectual Development : PHILIP E. VERNON : I	9
Higher Education, Class Influences in : C. T. SANDFORD, M. E. COUPER, and S. GRIFFIN	117
Subnormal Children : MAURICE CHAZAN	183
Morrisby Differential Test Battery, A Study of Some Contributions to Vocational Selection : M. K. L. EL-SHARKAWY and D. M. LEE	277
Perception, Developmental Features of Behaviour and : GERALD H. FISHER	223
Perception, the Structure of ; A Factorial Study : R. M. BEARD	69
Perceptual Organisation, levels of, and of Performance after time for Reflection : JOSEPHINE KLEIN	210
Personality, a New Scale for Measurement of in Children : SYBIL B., G. EYSENCK	60
Personality Ratings, of Pupils in Scottish Primary Schools, Teachers : A. MORRISON, D. MCINTYRE and J. SUTHERLAND	362
Programmes, a Comparison between Linear and Skip Branching : JAMES HARTLEY	306
Reading Attainment Test, the effect of Practice on a : C. NETLEY, S. RACHMAN and R. K. TURNER	320
Reading, Colour as an Aid to Visual Perception in Early : J. KENNETH JONES	1
Reading Failure and Unsettledness and Maladjustment : ROBIN GREGORY	21
Reading, The Psychology of Beginning : An Experiment with Two Methods : E. S. TENSUAN and F. B. DAVIS	63
Religious Attitude and Practices among Students, changes in during University Degree Courses : G. W. PILKINGTON, P. K. POPPLETON and G. ROBERTSHAW	127
Religious Story Data, the Application of Piaget's Schema of Operational Thinking to, by means of the Guttman Scalogram : R. J. GOLDMAN	150
Rigidity in Problem Solving, and Teaching Method : M. L. KELLMER-PRINGLE and I. R. MCKENZIE	158
SCIENCE PRACTICAL TESTS, the Evaluation of at the Secondary School Level : A. A. ABOUSEIF and D. M. LEE	50
Teachers, a Follow-Up of, Five Years After Completing their Training : S. WISEMAN and K. B. START	41
Teacher's Role, Teachers' and Parents' Conception of : F. MUSGROVE and P. H. TAYLOR	342
Teacher Trainee Selection, Personality Factors in : M. S. TARPEY	171
Valentine Reasoning Test, the Effect of Logic Instruction on : CHARLES F. ELTON	146

Research Notes :

Cattell and Eysenck, the Concepts of : C. J. ADCOCK	90
Extraversion, Neuroticism and the Effect of Praise or Blame : Y. RIM	381
Guessing Correction, the, in Vocabulary Tests : M. HAMMERTON	249
I.Q. Changes in Educationally Subnormal Children at Special School : ZENA A. STEIN and GREGORY STORES	379
J.M.P.I., Performance of Canadian Children on : A. J. CROPLEY and B. R. D'Aoust	378
Personality, Factorial Studies of and Their Bearing on the Work of the Teacher : C. BURT	368
Personality Test, a New for Boys : H. B. GIBSON	244
Reading Age, Multiple Regression Equations for Predicting from Chronological Age and WISC Verbal I.Q. : FAY FRANELLA and DAVID GERVER	86
Special Educational Treatment in the Ordinary Schools : H. G. ARMSTRONG	242

Critical Notices :

A Short History of British Psychology, 1840-1940, by L. S. Hearnshaw : R. S. PETERS	265
Education and Environment, by Stephen Wiseman : C. BURR	98
The Home and the School : A Study of Ability and Attainment in the Primary Schools, by J. W. B. Douglas : C. BURR	259
University Grants Committee Report of the Committee on University Teaching Methods : T. H. PEAR	385

Thesis Abstracts :

Logical Thinking in Normal and Sub-normal Children, the growth of : S. JACKSON	255
Reading Readiness and Reading Progress, the Relationship between : D. V. THACKRAY	252

Authors :

ABOUSEIF, A. A. and LEE, D. M. : The Evaluation of Certain Science Practical Tests at the Secondary School Level	41
ADCOCK, C. J. : A Comparison of the Concepts of Cattell and Eysenck	90
ARMSTRONG, H. G. : Special Educational Treatment in the Ordinary Schools	242
BEARD, R. M. : The Structure of Perception : A Factorial Study	210
BURT, C. : The Home and the School : A Study of Ability and Attainment in the Primary Schools, by J. W. B. Douglas	259
BURT, C. : Wiseman, Stephen (1964), Education and Environment.	98
BURT, C. : Factorial Studies of Personality and their Bearing on the Work of the Teacher	368
CHAZAN, M. : Factors Associated with Maladjustment in Educationally Subnormal Children	277
CROPLEY, A. J., and D'Aoust, B. R. : Performance of Canadian School Children on the J.M.P.I.	378
DALE, R. R. : Co-education : The Verdict of Experience, II—A Qualitative Approach	195
DOUGLAS, J. W. B., and ROSS, J. M. : The Effects of Absence on Primary School Performance	28
EL-SHARKAWY, M. K. L., and LEE, D. M. : A Study of Some Contributions of the Morrisby Differential Test Battery to Vocational Selection	223
ELTON, CHARLES F. : The Effect of Logic Instruction on the Valentine Reasoning Test	339
Eysenck, S. B. G. : A New Scale for Personality Measurement in Children	362
FRANSELLA, FAY, and GERVER, D. : Multiple Regression Equations for Predicting Reading Age from Chronological Age and WISC Verbal I.Q.	89
FERRON OSCAR : A Comparative Study of the Attitudes of West African Students and Teachers Towards the Modern Approach in Teaching	294
FISHER, GERALD H. : Development Features of Behaviour and Perception	69
GIBSON, H. B. : A New Personality Test for Boys	244
GOLDMAN, R. J. : The Application of Piaget's Schema of Operational Thinking to Religious Story Data by Means of the Guttman Scalogram	158
GREGORY, ROBIN E. : Unsettledness, Maladjustment and Reading Failure : A Village Study	63
HAMMERTON, M. : The Guessing Correction in Vocabulary Tests	249
HARTLEY, JAMES : Linear and Skip-Branching Programmes : A Comparison Study	320
JACKSON, S. : The Growth of Logical Thinking in Normal and Sub-Normal Children	255
JAMIESON, G. H. : Psychological Aspects of Craftsmanship in Pottery-Making at a Secondary School	179
JONES, J. KENNETH : Colour as an Aid to Visual Perception in Early Reading	21
KELLMER-PRINGLE, M. L. : Teaching Method and Rigidity in Problem Solving	50
KLEIN, JOSEPHINE : Levels of Perceptual Organization and of Performance After Time for Reflection	60
MORRISON, A., McIntyre, D. and SUTHERLAND, J. : Teachers' Personality Ratings of Pupils in Scottish Primary Schools	306
MUKHERJEE BISHWA NATH : Achievement Motivation and Goal-Setting Behaviour in the Classroom	286
MUSGROVE, F., and TAYLOR, P. H. : Teachers' and Parents' Conception of the Teacher's Role	171
NETLEY, C., RACHMAN, S., and TURNER, R. K. : The Effect of Practice on Performance in a Reading Attainment Test	1
PEAR, T. H. : University Grants Committee Report of the Committee on University Teaching Methods (1964)	385
PETERS, R. S. : A Short History of British Psychology, 1840-1940, by L. S. Hearnshaw	265

Authors—continued.

PILKINGTON, G. W., POPPLETON, P. K., and ROBERTSHAW, G. : Changes in Religious Attitude and Practices Among Students during University Degree Courses ..	150
RIM, Y : Extraversion, Neuroticism and the Effect of Praise or Blame ..	381
SANDFORD, C. T., COUPER, M. E., and GRIFFIN, S. : Class Influences in Higher Education ..	183
STEIN, ZENA A., and STORES, G. : I.Q. Changes in Educationally Subnormal Children at Special School ..	379
TARPEY, M. S. : Personality Factors in Teacher Trainee Selection ..	140
TENSUAN, E. S., and DAVIS, F. B. : The Psychology of Beginning Reading : An Experiment with Two Methods ..	127
THACKRAY, D. V. : The Relationship between Reading Readiness and Reading Progress ..	252
VERNON, PHILIP E. : Environmental Handicaps and Intellectual Development, Part I ..	9
Part II ..	117
WISEMAN, S., and START, K. B. : A Follow-up of Teachers Five Years after Completing Their Training ..	342
WISENTHAL, M. : Sex Differences in Attitudes and Attainment in Junior Schools ..	79
YAMAMOTO, KAORU : Effects of Restriction on Range and Test Unreliability on Correlation Between Measures of Intelligence and Creative Thinking ..	300
ZAHARAN, HAMED A. S. : A Study of Personality Differences Between Blind and Sighted Children ..	329



NOTES TO CONTRIBUTORS

1.—The Editor is glad to consider articles in the fields of educational and child psychology, whether written from the theoretical, clinical, experimental or psychometric standpoints. It is, however, desirable that statistical matter (formulae, tables, discussions of techniques, etc.) should be kept to a minimum, so that the *Journal* will be of interest to statistically untrained readers. Such concepts as the mean, percentiles, standard deviation, correlation, chi-squared, standard error, critical ratio, significance and reliability, and general or group factors, can be employed in the text. But more complex concepts and methods should be explained or confined to footnotes or an appendix, or published elsewhere.

Articles which have been, or are to be, published elsewhere cannot be accepted.

2.—*Length.* The usual range of length is from 2,500 to 4,500 words, and only very exceptionally will articles of more than 6,000 words be published. Short research reports of 500—1,000 words, such as outlines of theses, are welcomed. These should normally be submitted through the Head of the Department in which the research was carried out.

3.—*Arrangement.* Articles should be preceded by a summary or abstract of between 100 and 300 words, embodying the main conclusions. This should be typed directly after the title and author's name and place of work. A contents analysis, or list of headings, need *not* be included, but successive sections of the article should receive numbered headings, e.g.:

- I. INTRODUCTION.
- II. OUTLINE OF PREVIOUS RESEARCH.
- III. PLAN OF THE INVESTIGATION.
 - (i) *Subjects and materials.*
 - (ii) *Experimental Procedure.*
- IV. STATEMENT OF RESULTS.
- V. DISCUSSION AND CONCLUSIONS.
- VI. REFERENCES.

4.—*Typing.* All papers must be typed (double or 1½-spaced) on one side of quarto or foolscap paper. There should be margins of at least one inch all round. Authors should consult articles in current numbers of the *Journal*, and copy their conventions in respect of capitalising the title and main section headings, italicising author's place of work and sub-section headings, indentation of paragraphs, the form of references and tables, etc.

5.—*Footnotes* should be few in number, and should be indicated in the text by using the signs *, †, ‡, §.

6.—*References* should normally be indicated by date following the author's name in the text, e.g., Spearman (1904), or (Spearman, 1904, p. 160), or—where there are two or more publications in the same year—Burt (1939a). However, it is occasionally more convenient to number the references throughout, and to type numbers in the text above the line, e.g., Spearman.⁴ When there are two or three references only, these may be placed in footnotes in the text, instead of in a separate section at the end.

The list of references at the end should always be in alphabetical order of author's names, and—for any one author—by order of date. The following standard form of references should be observed:

DEARBORN, W. F., and ROTHNEY, J. W. M. (1941). *Predicting the Child's Development*. Cambridge, Mass.: Sci.-Art.

(Note that Ltd., etc., after publisher's names, also the numbers of pages in the book, can be omitted.)
BURT, C. (1939a). The relations of educational abilities. *Brit. J. Educ. Psychol.*, 9, 55-71.

(Note the absence of inverted commas and of capital initials in the titles of articles. The number of the volume is in arabic, not roman numerals. The part of the journal is not given, but page numbers are included in full.)

The World List of abbreviations (as given in the December issues of *Psychological Abstracts*) is used for journals.

7.—*Tables and figures* are costly to print, and should be used only when they save space and/or are essential for effective presentation. The same material should not be presented in a table and a graph. Each table or figure should be numbered (in arabic figures), on a separate sheet, and its approximate position indicated in the typescript. Drawings should be in Indian ink on heavy, unruled paper, the lettering on graphs being in pencil. Lines in tables should be ruled in pencil. Each table or graph should have a short title, sufficient for understanding without searching in the text.

8.—*Corrections.* The Editor cannot consider for publication articles which are seriously deficient in presentation. The typescript submitted (after preliminary consultation, if desired) should represent the final form in which the author wishes the paper to appear. Since changes in proof are costly, out of all proportion to the original setting, authors must be charged for any changes, insertions or deletions other than printer's errors).

9.—*Reprints.* Authors receive fifty copies of their papers free. Extra copies, in multiples of fifty, may be had at cost price if the order is given when the proof is returned.

GOVERNMENT SOCIAL SURVEY

In order to carry out expanding programmes of research in many fields additions to research and technical staff are needed. Applications are invited for the posts of Principal and Senior Survey Research Officers.

The Social Survey is the Government's central survey research organisation which works for all departments, Royal Commissions and other public Committees. Current studies include work in Housing, Delinquency, Manpower, Education, Social Economics and Consumer Expenditure, Transport and many aspects of Social Administration and Welfare Services. A wide range of methods is employed involving both factual study as well as opinion and attitude research, and emphasis is laid on the development of appropriate methodology. In the Social Survey research officers work with specialist technicians in sampling, interviewing, coding and computing but have full responsibility for the design and reporting of their project.

Most Social Survey reports are published and the contribution of papers on research methodology to professional journals is encouraged.

All applications should be qualified in one of the Social Sciences, Statistics or Mathematics or have had relevant experience in social, economic or operation research.

PRINCIPAL SURVEY RESEARCH OFFICERS should have shown themselves capable of organising a research programme, have carried out studies showing advanced technical skill and be capable of supervising the work of the other research officers. Salary rising to £2,983.

SENIOR SURVEY RESEARCH OFFICERS should be able to produce several complete reports for which they have had major responsibility. Salary rising to £2,414.

Applicants should list all previous work and be able to produce reports of completed studies.

These posts can lead to superannuated established positions and there are prospects of promotion to higher levels ranging at present up to £4,385 p.a.

Write for application form to Manager (PE 2630/UJ.) Ministry of Labour, Professional and Executive Register, Atlantic House, Farringdon Street, London, E.C.4.

Interested research workers who would like further information about the work or conditions may contact the Director, Government Social Survey, Telephone CITY 5744 Ext. 291.